Cognitive Psychology 157 (2025) 101715

Contents lists available at ScienceDirect

Cognitive Psychology

journal homepage: www.elsevier.com/locate/cogpsych

Word order effects in sentence reading

Petar Atanasov [a,*], Simon P. Liversedge [a], Federica Degno [b]

[a] School of Psychology and Humanities, University of Central Lancashire, Preston PR1 2HE United Kingdom
[b] Department of Psychology, Bournemouth University, Poole BH12 5BB United Kingdom

ARTICLE INFO

Keywords:
Eye Movements
Reading
Word Transpositions
Grammaticality Decisions

ABSTRACT

The SEAM model (Rabe et al., 2024) and the OB1-Reader model (Snell, van Leipsig, et al., 2018) suggest that readers lexically process words in parallel, with the OB1 model further specifying that those words are formed into a sentence-level representation irrespective of their order of presentation. The serial model, E-Z Reader (Reichle, 2011), in contrast, stipulates that words are identified serially and sequentially. The current eye tracking experiment investigated whether, how frequently, and how rapidly readers detect sentential anomalies arising from word transpositions and ungrammatical sentence final words. We also assessed the consequences in the eye movement record of processing such transpositions and ungrammaticalities to evaluate theoretical claims extrapolated from different eye movement models. This was done via target word pair (transposed vs. non-transposed) and a final word grammaticality (grammatical vs. ungrammatical) experimental manipulations. Readers were better at judging the grammaticality of sentences containing both a word transposition and an ungrammatical final word than those with solely a word transposition. Critically, transposed words caused significant disruption to reading, but not prior to readers fixating the first word of the transposed word pair. Furthermore, an ungrammatical sentence-final word attracted readers' fixations and caused increased re-reading in the absence of a word transposition compared to when it was preceded by a transposed word pair. Together the results show the importance of canonical word order for natural undisrupted reading and question claims for parallel lexical identification in relation to eye movement control during reading.

1. Introduction

There exists a large body of evidence suggesting that processing of letters within a word occurs in a parallel fashion, and the encoding of letter positions is subject to a level of uncertainty (e.g., Gomez et al., 2008). This body of evidence is, to some extent, based on letter transposition effects, whereby nonwords created by swapping the order of two letters within the original word (e.g., *dcotor* from *doctor*) are responded to, or read faster, compared to nonwords created by substituting two of the letters within the original word (e.g., *dmator*; see Johnson et al., 2007; Pagan et al., 2021).

Higher order transpositions at the level of morphemes and words have remained relatively under-investigated, with only a single study on word transpositions prior to 2018 (i.e., Rayner et al., 2013). Rayner and colleagues investigated the influence of word transpositions presented in parafoveal vision via the boundary paradigm (Rayner, 1975), and compared reading times when participants received an identical preview (e.g., *white walls* –> *white walls*), a transposed-word preview (e.g., *walls white* –> *white walls*; that

* Corresponding author at: University of Central Lancashire, Preston PR1 2HE, United Kingdom.
E-mail address: PFAtanasov@uclan.ac.uk (P. Atanasov).

https://doi.org/10.1016/j.cogpsych.2025.101715
Received 7 May 2024; Received in revised form 3 December 2024; Accepted 14 January 2025
Available online 16 February 2025

did not constitute a grammaticality violation), or a substituted word preview (e.g., *vodka clubs* −> *white walls*). Their findings showed that both invalid preview conditions caused disruption with longer reading times on the post-boundary word(s). However, that cost was graded such that when the preview consisted of substitute words, reading times were longer than when the preview consisted of a transposed word pair. Rayner et al. interpreted this pattern as evidence that disruption to processing increases as the linguistic overlap between preview and targets decreases.

While Rayner et al. (2013) studied the effect of word transposition in parafoveal vision during natural reading, work on word transpositions in the past six years has instead been focused primarily on the detectability of transposed words that create a grammaticality violation (e.g., Mirault et al., 2018). This line of experimentation has prompted researchers to consider whether positional encoding and processing of words within sentences or word strings might be similar to positional encoding and processing of letters (Dufour et al., 2022; Huang & Staub 2021a, 2021b, 2022, 2023; Hussain & White, 2023; Liu et al., 2020, 2021, 2022; Milledge et al., 2023; Mirault et al., 2018, 2020; Mirault, DeClerk, et al., 2022, 2023; Mirault, Leflaec et al., 2022; Mirault, Vandendaele, et al., 2022, 2023; Pegado & Grainger, 2019, 2020a, b; Snell & Grainger, 2019a, 2019b; Snell & Melo, 2024; Snell et al., 2023; Spinelli et al., 2024; Tiffin-Richards, 2024; Wen et al., 2019, 2021a, 2021b, 2022, 2024).

A striking finding from Mirault et al. (2018, 2020) is that readers sometimes fail to notice word transpositions within a sentence (e.g., *The white <u>was cat</u> big*) if a plausible representation of that sentence can be formed (e.g., when the words can be rearranged to be meaningful and syntactically correct, as in *The white cat was big*). This failure to detect a word transposition is particularly pronounced when compared to a situation where such a representation is not possible (e.g., when the final word in a string precludes syntactic legality of that string, as in *The white <u>was cat slowly</u>*). The *Transposed-Word* effect has also been found to be greater (i.e., slower responses and more detection failures) when the transposed word pair is internal (e.g., *The can man run*) compared to external (e.g., *Run man can the*; Snell & Grainger, 2019b).

The *Transposed-Word* effect has been replicated in different languages (in Chinese: Liu et al., 2020, 2021, 2022; Dutch: Snell & Melo, 2024; German: Tiffin-Richards, 2024; English: Hossain & White, 2023; Huang & Staub, 2021a, 2022, 2023; Milledge et al., 2023; Spinelli et al., 2024). Effects of word transpositions with comparable magnitudes have been observed in both speeded (e.g., Liu et al., 2020, 2021; Mirault et al., 2018, 2020) and unspeeded (e.g., Liu et al., 2021) grammaticality decisions, suggesting that the effect is robust even when participants do not need to sacrifice accuracy to perform the task as quickly as possible (a speed-accuracy trade-off). Beyond grammaticality decisions, word transpositions have been shown to affect change detection (e.g., Pegado & Grainger, 2020) such that participants are worse at detecting a change that was created by transposing two words than by replacing them, suggesting that word transpositions may be difficult to be detected regardless of the specific task constraints.

The recent OB1-Reader model (Snell, van Leipsig, et al., 2018) aims to explain *Transposed-Word* effects. The model specifies that multiple words are lexically processed in parallel during natural reading. According to OB1-Reader, semantic information is not immediately integrated across words within a sentence (Snell & Grainger, 2019a). Instead, when words are identified in parallel, semantic information associated with each of those words is tied to the particular location assigned to that word in the sentence frame via a spatiotopic sentence-level representation. The individual semantic meanings of the words are then integrated with sentential context at a later, integration, stage of processing. In this way, the model allows for parallel lexical identification of words whilst clearly specifying that the semantic characteristics of words that appear later in the sentence will not influence processing of words that appear earlier in the sentence.

According to Snell, van Leipsig, et al. (2018), readers form a sentence-level spatiotopic representation upon first fixating a sentence, based on word length (although see Snell et al., 2023; Wen et al., 2024) and syntactic expectations. Based on these assumptions, then, it is quite possible that word order determined by the reader might not match the actual order of the words that appear within a sentence. In this way, the transposed-word phenomenon, that is, failure to detect transposed words in otherwise grammatically correct sentences, occurs due to successful location tagging of words in relation to their correct positions in a sentence. Furthermore, word position coding is handled by the spatiotopic sentence-level mechanism in working memory such that the assignment of words to positions in a sentence does not affect oculomotor control. Therefore, any disruption associated with detecting a word transposition (i.e., detecting the ungrammaticality) should be observable in global measures such as response times and accuracy on the grammaticality decision task. Conversely, no disruption should be observable in (local) eye-movement measures on the critical transposed words. Currently, OB1 Reader is the only computational model of eye movement control during reading which directly aims to explain both eye movement patterns and detection rates associated with ungrammatical word transpositions embedded in sentences.

The autonomous Saccade generation With Inhibition by Foveal Targets (i.e., SWIFT: Engbert et al., 2002; 2005) is a parallel gradient model of eye movement control that chronologically preceded the OB1 Reader model (though note that the two models were developed quite independently by different groups of researchers). According to SWIFT, multiple words can be lexically identified in parallel during sentence reading. Crucially, for the present paper, SWIFT does not model how words are integrated within context. Instead, the model only focuses on visual and lexical processing as well as saccade programming. In SWIFT, there is a lexical processing stage, wherein a word is activated based on visual input (and in relation to word length, lexical frequency and predictability characteristics). When the activation level of a representation corresponding to a word reaches a required threshold, lexical identification is assumed to be completed. Given that words in a sentence are identified in parallel, acording to SWIFT, it is possible that transposed words that appear out of order in the sentence might be processed quite successfully, although this explanation would require that those words would be lexically identified out of order on any occasion that a reader fails to detect a transposition. Although this might be possible in principle, according to the SWIFT model, parafoveal words further from the point of fixation most often receive less activation than parafoveal words closer to fixation due to visual constraints. This means that words further into the parafovea are usually identified more slowly than those closer or in foveal vision, and therefore, an explanation of effective transposed word processing based on consistent out of order lexical identification seems unlikely.

While the SWIFT model may not offer a direct account of word transposition effects, a recent related, successor, model – SEAM (Rabe et al., 2024) has been put forward, which has the capability of predicting when the influence of a grammatically illegal transposed word pair might first have an influence on eye movements during reading. The SEAM model combines the visual and lexical processing, as well as the saccade planning assumptions of SWIFT, with a set of syntactic integration assumptions associated with the Activation-Based model (Lewis & Vasishth, 2005). The Activation-Based model postulates that words are integrated into context on the basis of a set of rules held in procedural memory. For example, in the sentence *The robber that the policeman in the patrol car chased escaped*, as soon as the words *robber* and *policeman* are lexically identified, they are stored in working memory since, syntactically, both are associated with constituents that appear later in the sentence (in this sentence, the verbs *chased* and *escaped*). As readers process the sentence, when they identify the word *chased* downstream in the sentence, a retrieval process is initiated to retrieve from working memory the appropriate noun that is dependently associated with it (*policeman*). Notably, memory retrieval is subject to a degree of noise meaning that in some instances it is possible that instead of correctly identifying *policeman* as the subject of *chased*, readers may mistakenly bind *chased* with *robber*. Subsequent to the retrieval stage, there is a post-retrieval stage of processing where activations associated with words in working memory return steadily to zero. The main point to note, here, is that in SEAM, there are two additional stages of processing that have been appended beyond those associated with SWIFT, that are associated with the incorporation of new words into the representation of sentential meaning. Based on these new stages of processing, SEAM offers an account of where, and when, regressive saccades will be made during reading that is particularly helpful in considering eye movements and processing associated with syntactically ambiguous garden path sentences and sentences in which there may be ambiguity in the formation of coreference relations.

In contrast to parallel models, the E-Z Reader Model (Reichle et al., 1998; Reichle, 2011) is a Serial Attention Shift model. According to this model, words are processed in stages, serially and sequentially, whilst a process of integration runs in the background. Recognition of a word during reading occurs initially via a preattentive visual stage of processing whereby the visual characteristics of the upcoming word are processed. This is followed by an early lexical stage of processing (a familiarity check) wherein readers discern the letter make-up, that is, the orthography and potentially the phonology of the word. Next, during a later lexical stage, full lexical access of the word occurs, and this takes place as the word is directly fixated. Most of the time, the word to the right of the currently fixated word, in the parafovea, can only be pre-processed up to and including the early lexical stage before it is fixated. Furthermore, similar to SWIFT, E-Z Reader does not directly model how words are assigned to positions in text, and instead, words are simply appended to the end of the word sequence once they are identified.

Before considering the models further, there are several points of note. First, SEAM and E-Z Reader are models of eye movement control that seek to account for behaviour during normal reading of meaningful sentences (indeed, this is a characteristic of all models of eye movement control during reading other than the OB1 model). SEAM and E-Z Reader are not models of the eye movement behaviour that occurs after readers have identified that a sentence is ungrammatical. Thus, it is unreasonable to anticipate that the SEAM and E-Z Reader models deliver predictions about patterns of eye movements after readers encounter a word sequence indicating a transposition or an ungrammaticality. Presumably, after encountering a transposed or ungrammatical word, readers would experience a catastrophic failure in processing due to that word being illegal in respect of the current syntactic analysis under construction. And in such a situation it is likely that the reader would form a decision that the sentence is ungrammatical and consequently make a button press response accordingly. To reiterate, eye movements under such circumstances do not reflect natural reading, but instead reflect a secondary task (grammaticality decision formation) and are therefore processes that SEAM and E-Z Reader were not designed to explain.

Given these points, one may question the relevance of these models of eye movement control in reading to the current experiment. In our view, these models are very relevant to one of the theoretical issues we wished to explore in our experiment, namely, how rapidly readers detect sentential anomalies arising from word transpositions and ungrammatical sentence final words. We say this for two reasons. First, we assume that in order for readers to process sentences and make a decision as to whether or not they are grammatical, then they must actually read those sentences. To us, it is unclear how a grammaticality decision for a sentence could be made without an individual reading that sentence to assess whether it is grammatically well formed. Thus, up to the point that a reader detectes a transposed or ungrammatical word, we assume that they woudl be reading the sentence normally. Second, we assume that any catastrophic failure in processing that might occur upon detection of a transposed or ungrammatical word would cause increased fixation durations and increased (first pass) reading times at the point that the transposed or ungrammatical word was first detected. If this is the case, and our understanding of the models is correct, then the SEAM and E-Z Reader models make diverging predictions regarding exactly when disruption might first appear in the eye movement record upon detection of an illegality. To us, this is a very important point as it speaks both to when readers lexically process a word and integrate it with sentential context and to whether readers lexically identify words serially and sequentially (as specified by E-Z Reader) or in parallel (as specified by SEAM). If readers lexically identify words serially and sequentially, then we might anticipate disruption by a word that causes catastrophic failure in processing to occur only when the eyes fixate that word. Alternatively, if readers lexically identify words in parallel, then such disruption should occur on fixations prior to the transposed or ungrammatical word (since, at least on some occasions, words will be identified as being illegal prior to their fixation). Thus, in our view, the stipulations of the SEAM and E-Z Reader models are critical in respect of exactly when readers first exhibit disruption to processing due to transposed and ungrammatical words. Furthermore, beyond the predictions of these models, since the OB1 Reader model was designed specifically to account for processing of transposed words, it is possible to develop a much more comprehensive set of predictions for this model in relation to whether and when illegalities might first be identified, as well as how processing might proceed beyond that point. Accordingly, we provide detailed predictions from each of the models below.

It is also necessary to briefly explain why we chose to adopt a grammaticality decision task in our experiment. First, most eye movement experiments investigating transposed word effects in reading have required participants to make grammaticality decisions about sentences with, and without, transposed and ungrammatical words. Thus, to allow for comparability of our findings with other published studies, we opted for the same experimental task. The second reason we adopted a grammaticality decision was through necessity in respect of our expeirmental objectives. Transposed word studies require a grammaticality decision in order that researchers know whether a transposition was, or was not, detected. When readers fail to detect a transposition, reading and eye movement behaviour is indistinguishable from reading and eye movements for normal grammatical sentences (Huang & Staub, 2021a). And, as noted earlier, according to the OB1 Reader model, when readers process a transposed word pair, those words are ordered appropriately in a secondary "interpretation" stage of processing. This would mean that the sentence would be processed by the reader normally regardless of whether it did, or did not, contain a word transposition, and regardless of whether the reader did, or did not, detect a transposition. For this reason, to test the OB1 Reader model, it is critical to establish whether the reader judged the sentence as grammatical or ungrammatical. Only on this basis (according to the model) would it be possible to know whether a transposition was detected. When a transposition was detected, the reader should provide an ungrammatical response, and when the transposition was not detected, the reader should provide a grammatical response. Critically, however, under both circumstances, the eye movement behaviour, according to the OB1 model, should not differ. Thus, to evaluate the OB1 model's stipulations regarding patterns of eye movements in relation to transposed words in sentences, it was vital to assess both eye movement behaviour and grammaticality decisions.

Although a number of studies have been taken to support the OB1-Reader model (e.g., Mirault et al., 2018, 2020), there is also some research that has challenged this account (Hossain & White, 2023; Huang & Staub, 2021a, 2022, 2023; Liu et al., 2020, 2021, 2022; Milledge et al., 2023; Spinelli et al., 2024; though see also Mirault, Vandendaele et al., 2022; Snell & Melo, 2024). Liu et al. (2022) included word transpositions in grammatical and ungrammatical base sentences, such that the grammatical base sentences contained a final syntactically legal word while the ungrammatical base sentences contained a final syntactically illegal word. The authors found a word transposition effect regardless of whether the sentences were presented one word at a time in the middle of the screen, to encourage serial lexical encoding, or presented all at once on the screen, to encourage parallel processing. Although, notably, the *Transposed-Word* effect in the serial mode presentation was weaker and only reliably obtained for accuracy and not for RTs, while the effect was observed for both accuracy and RTs in the parallel mode. This led Liu et al. (2022) to conclude that transposition effects do not necessarily indicate that word processing needs to occur in parallel (see also Huang & Staub, 2023). Instead, they argued that the failure to notice words out of order might be explained in line with *noisy-channel* theories (e.g., Gibson et al., 2013). That is, readers might initially detect the word transposition, and thus, the sentence ungrammaticality, but due to top-down predictive mechanisms, such as syntactic and semantic expectations, they might still be able to form a plausible interpretation and adopt this, thereby overriding the syntactic violation (see also Liu et al., 2020). Subsequently, Huang and Staub (2023) found a similar pattern of effects for serial versus parallel presentations and argued parallel lexical processing is not a requirement for *Transposed-Word* effects.

Mirault, Vandendaele, et al. (2022) observed the same pattern of effects as found by Liu et al. (2022) and Huang and Staub (2023) and argued that the lack of a *Transposed-Word* effect on RTs when sentences are presented serially, word by word, provides evidence in favor of the OB1 Reader account since syntactic expectations still drive word position coding even in the absence of a defined spatiotopic sentence-level map. Recently, Snell and Melo (2024) also failed to find any *Transposed-Word* effects when presenting sentences word by word in Dutch, which they argued is further evidence that word position coding is based on word length cues and syntactic expectations and these exert a weaker influence in serial presentation scenarios than in parallel presentation scenarios.

Hossain and White (2023), recently provided evidence, however, that *Transposed-Word* effects can be reliably obtained for RTs and accuracy when presenting words serially when the presentation speed is matched to that of the reader. In addition, Milledge et al. (2023) observed a robust *Transposed-Word* effect on accuracy when using a serial presentation mode but with words presented in their correct sentential positions rather than in the same central location. Spinelli et al. (2024) further tested the existence of *Transposed-Word* effects in sentences presented naturally (Experiment 1) and serially, word-by-word in the same central location or with the correct location of each word preserved (Experiment 2). Crucially, the authors included an additional condition which included an ungrammatical sentence final word only. This enabled the authors to compare the impact of different types of grammaticalilty violations on grammaticality decisions. In their first experiment, the authors found that accuracy and response times were similar for sentences only containing word transpositions compared with sentences only containing a final ungrammatical word. In both experiments, participants were better and faster at judging sentences containing a word transposition and a final ungrammatical word than sentences containing a word transposition only, replicating the *Transposed-Word* effect. However, this Transposed-Word effect was diminished in their second experiment in comparison to the effect observed in the first experiment, where all the words in a sentence were presented simultaneously. Interestingly, the authors argued that the results across both experiments did not fit within the current specification of the OB1 Reader model and suggested instead that readers may rely on temporal cues (rather than spatial cues) associated with each word specifying when its processing started.

Whether *noisy-channel* theories (e.g., Gibson et al., 2013) are able to explain the transposition effects was directly examined by Huang and Staub (2021a) via two eye tracking experiments. The authors hypothesised that if those theories are correct, the data should show an effect of word transposition in early eye movement measures, even when readers fail to notice that transposition, and that this effect should be reduced in later measures once readers override the grammatical violation in favor of a plausible sentence representation. In their first experiment, participants engaged in a grammaticality decision task. In their second experiment, participants were asked to read each sentence and then answer a question concerning its grammaticality, or instead a comprehension question pertaining to its meaning. Participants were not forewarned of the type of question they would be required to answer on any particular trial. Regardless of the type of questions participants were required to answer in the two experiments, Huang and Staub found that there were more grammaticality decision errors when judging sentences with transposed words compared to grammatically correct sentences, thus confirming that sometimes readers do fail to notice words out of order. Moreover, the eye movement data showed disruption at the point of ungrammaticality only when participants detected the violation, but not when they failed to detect that the words were out of order. Early reading times were comparable at the point of ungrammaticality between grammatically correct sentences and sentences with a transposition that was not noticed. Note, however, for this latter condition, the eye movement record did show some disruption to reading on later measures. Specifically, Huang and Staub showed inflated rates of regressions to the text preceding the ungrammaticality and increased total viewing times for the word preceding the ungrammaticality for Experiment 1, and increased regression-out rates for the word preceding the ungrammaticality for Experiment 2. These findings were interpreted by the authors as evidence that some form of top-down mechanism must play a role in word processing, as readers do, in fact, sometimes fail to notice the word transpositions. However, the authors argued that any such top-down mechanism must operate much earlier than *noisy-channel* accounts would predict, as there were no effects of word transposition in the early eye movement measures when readers failed to notice the violation. Thus, Huang and Staub (2021a) proposed that word transpositions can be explained by a serial processing account (e.g., the E-Z Reader model, Reichle et al., 1998; Reichle, 2011) in which semantic and syntactic integration can proceed in a non-incremental fashion, with integration of word n + 1 at times taking place before integration of word n is completed (see Huang & Staub, 2021b for a detailed discussion).

Although more research is needed to establish what causes readers to fail to notice word transpositions, two additional findings from Huang and Staub (2021a) seem to contradict the OB1-Reader explanations for this effect. In their two experiments, Huang and Staub transposed target word pairs in the third and fourth position of the sentence, such that the target words were either both short, or one of them was long (short: 2–4 letters, long: 8–11 letters). According to OB1-Reader, for a word to assume a particular sentential position, it must match the perceived length of the word expected in that position (based on parafoveal and peripheral processing). If this is correct, then it follows that when transposed words are of different lengths, particularly when that length difference is substantial, readers should detect the transposition. However, in the Huang and Staub (2021a) experiments, readers failed to detect a transposition even when the target words were of very different length. Additionally, in Experiment 1, Huang and Staub found that the transposition effect on accuracy was modulated by the likelihood of skipping either of the words in the transposed word pair. When participants skipped either of the transposed words, they were more likely to judge the sentence as grammatically correct showing that changes in eye movement behavior in and around transposed words modulate ungrammaticality detection. This should not be the case according to the OB1 Reader model.

In a study by Mirault et al. (2020), where participants' eye movements were recorded via a virtual reality setup, transposed target word pairs were embedded in sentences with either a syntactically legal, or an illegal, final word. Eye movements were compared for sentences containing a transposition and an illegal final word against those for sentences with a transposition and a legal final word. Mirault et al. (2020) found lower accuracy for sentences with a transposition and a legal final word than for sentences with a transposition and an illegal final word. However, while there were no effects of word transposition on eye movements on the target word pair, there were longer total sentence reading times for the sentences with a transposition and a legal final word compared to the sentences with a transposition and an illegal final word. It is likely that the lack of effects for local eye movement measures for the two transposed words were due to the authors comparing reading times for transposed words in sentences with a legal versus illegal final word. That is to say, the lack of a difference was likely due to the fact that the reading times were compared for word pairs that were transposed in both experimental conditions.

2. The present experiment

The present experiment aimed to expand on previous research on word transposition effects considering existing computational models of eye movement control during reading, namely E-Z Reader (Reichle et al., 1998; Reichle, 2011), SEAM (Rabe et al., 2024), and OB1-Reader (Snell, van Leipsig, et al., 2018). Further, since Mirault et al. (2018; 2020) only directly compared sentences containing a transposition with a syntactically legal versus illegal final word, we sought to disentangle effects arising due to word transpositions and effects arising due to the syntactic legality of the final word of the sentence. This is important because the grammaticality decision task is aimed at exploring how well participants are able to detect ungrammaticality in sentences. Note, though, that the task does not require participants to quantify how many ungrammaticalities they have detected, just whether they have, or have not, detected one. It is possible that the magnitude of *Transposed-Word* effects may be increased because having two ungrammaticalities in a sentence facilitates ungrammaticality detection, while a single ungrammaticality may be less easy to detect. Hence, adding a new condition in the current experiment, in which readers are presented with sentences containing a single violation to syntax that is not created by transposing two adjacent words, might allow us to determine whether the magnitude of the *Transposed-Word* effect is decreased for sentences containing two, versus one, violation to syntax.

We, therefore, manipulated final word grammaticality independently from word transposition, such that the final word could be grammatical or ungrammatical, and the third and fourth words of each sentence were presented in their correct order or transposed. We also ensured that the first point of ungrammaticality was always the third word of each sentence (i.e., the first word of the transposed pair). Furthermore, since in Mirault et al. (2018; 2020) word frequency and length were not controlled, and in Huang and Staub (2021a) they were manipulated, we chose to match our target words on both characteristics. Matching target words on frequency and length should maximise the chances of participants failing to detect the word transposition, and thus allow us to examine whether, even under these circumstances, eye movements might show some disruption.

Based on Mirault et al. (2018, 2020) and Huang and Staub (2021a Experiment 1), we predicted that readers would be significantly worse at judging the grammaticality of sentences containing only a transposition compared to sentences with a transposition and an ungrammatical final word. We also anticipated that the proportion of correct responses would be reduced for sentences containing only a word transposition than for correct sentences, though of course, participants would make *no* responses for the former and *yes* responses for the latter. We had two diverging expectations regarding performance on sentences containing only a syntactically illegal final word in comparison to sentences containing two violations to syntax. First, if the *Transposed-Word* effect is driven by the ease with which an ungrammaticality is detected in the presence of two compared to one syntactic violation, then participants should be better (i.e., more accurate and faster) in detecting the ungrammaticality in sentences with a word transposition and a final syntactically illegal word compared with sentences containing a final syntactically illegal word only. These predictions are supported by evidence from grammaticality decision studies in the auditory domain (e.g., Devescovi et al., 1997; Lu et al., 2000) as well as a recent grammaticality decision study on written text (Spinelli et al., 2024). If, on the other hand, the OB1 Reader account of processing is correct, then performance for sentences containing a final ungrammatical word only, should be comparable to performance for sentences containing two syntactic violations since rearranging incorrect word order would result in a syntactic violation associated with the ungrammatical final word in both conditions. In line with this reasoning, performance for sentences containing a final ungrammatical word only should be increased in comparison to sentences containing a transposition only, since the former cannot be resolved into a grammatical sentence, while the latter, can.

With respect to eye movements, our predictions were driven both by existing research on *Transposed-Word* effects, as well as by the different computational accounts discussed above. According to the E-Z Reader model, there should be a significant effect of transposition starting at the first point of ungrammaticality in the sentence (i.e., at Target Word 1) when the two target words were presented out of order, with longer first-pass reading times when words were transposed than when they were presented in the correct order. E-Z Reader would also predict a significant effect of the grammatical legality of the final word at that word, as a serial perspective predicts words will be lexically identifed serially and sequentially.

If readers process words in parallel, as per SEAM (Rabe et al., 2024), and assuming that transposed and ungrammatical words cause catastrophic failure in processing with quite immediate disruption to eye movements, then in line with a processing span of two words to the right of the fixated word, there should be an effect of transposition possibly two or one words prior to Target Word 1. Following a similar rationale, when the final word in the sentence is ungrammatical relative to when it is grammatical, it should also be possible to observe effects up to two words earlier in the sentence. Thus, at Target Word 1 and Target Word 2, if we compare sentences without a transposition and with a grammatical final word against those for sentences without a transposition but with an ungrammatical final word, we might see longer first-pass reading times but lower probability of regressions out and shorter go-past time in the latter than the former condition. Regardless, SEAM would predict that an effect of transposition should be detected at least on Target Word 1, and an effect of final word grammaticality on the Post-Target, both in terms of first-pass reading times (as for E-Z Reader), and in terms of probability of regressions out and go-past time.

Finally, if readers process words in parallel, as per the OB1 Reader model (Snell, van Leipsig, et al., 2018), we should only observe an effect of final word grammaticality, and this should only occur on the fifth word (the Post-Target word) in the sentence, with reading times being longer for ungrammatical than grammatical final words. Word transpositions should have no effect on early eye movement measures associated with any word within the sentence. That is to say, sentences with, and without, transpositions should have comparable reading times. Furthermore, assuming the OB1 model is correct, we should not observe a significant interaction between word transposition and final word grammaticality on early reading times. To be explicit, we should observe comparable early reading times on the Post-Target word for grammatical sentences and sentences containing a word transposition only. However, early reading times should be comparably inflated for both sentences containing a transposition and a final ungrammatical word and sentences containing a final ungrammatical word only.

3. Data availability

Stimuli information, data, analyses scripts, trimming procedures for the generalised linear mixed effects models, software information, and tables with random effects estimates from the generalised linear mixed effects models for all measures are available at: https://osf.io/uybpv/?view_only = 371d7b96eac94be481a5378af74d1d10.

4. Methods

4.1. Power analysis

To determine our target sample size, we conducted a power analysis via the PANGEA software by Westfall (2016; https://jakewestfall.org/pangea/) based on Mirault et al. (2020) and Cohen's d_z of 0.392 for sentence total viewing times, computed for the comparison between sentences containing a word transposition only versus sentences containing both a word transposition and a syntactically illegal final word (i.e., the only comparison they ran in their statistical analyses). The analysis indicated that with an effect of this size and 108 stimuli, 40 participants would be needed to obtain a power ≥ 0.9.

4.2. Participants

We tested a total of 45 native English speakers with no known reading impairments and normal or corrected-to-normal vision from the student and staff community of the University of Central Lancashire. Participants were recruited via SONA and social media posts to participate in the eye-tracking experiment. All participants received £6 in Amazon vouchers or 10 course credits to take part. Five participants were excluded from data analysis (four due to technical issues, and one due to an accuracy rate lower than 80 %), and the data of 40 participants (*M* = 20.8 years, *SD* = 2.5, Female = 34; Age Range = 18–30 years) was included in the final analyses.

4.3. Design

The study was conducted as a 2(transposition: not transposed vs. transposed) × 2(final word grammaticality: grammatical vs. ungrammatical) repeated-measures Latin-square design. With this design, each participant saw only one of the four versions of each sentence, and each sentence appeared an equal number of times across participants and its four versions.

4.4. Apparatus

Viewing was binocular but only participants' right eye movements were recorded using an SR Research Eyelink 1000 Plus system with a sampling rate of 1000 Hz. Participants were seated 70 cm from of an LCD monitor with 1920 by 1080 FHD resolution and 240 Hz refresh rate. Stimuli were presented with a horizontal offset of 240 pixels from the center of the monitor, in black on a grey background, and written in monospaced Courier font size 24, with 2.3 letters subtending 1° of visual angle. The experiment was designed and presented via Experiment Builder v2.3.38 (SR Research).

4.5. Materials

One hundred and forty sentences were initially created for the study. Each sentence was comprised of five words which constituted the five regions of interest (ROIs) used for statistical analyses: first word; pre-target word (always the second word in the sentence); target word 1 (in the third or fourth sentence position dependent on the experimental condition); target word 2 (in the fourth or third sentence position dependent on the experimental condition) and post-target word (always in the fifth and final sentence position). The order of the two target words (target word 1 and target word 2) was manipulated such that they could be presented in the non-transposed, correct order, or have their order swapped, such that the sentence contained a word transposition and became ungrammatical at the third word position (see Fig. 1). Additionally, the final word (post-target) was chosen such that it could either fit the sentential context created by the first four words (final grammatical word) or be an ungrammatical continuation (final ungrammatical word).

To maximise the likelihood of obtaining at least one fixation on each word of the sentence, the pre-target, target word 1; target word 2 and post-target were at minimum four letters long. The pre-target word was between 4 and 7 letters long (*M* = 5.44, *SD* = 0.92) and always a different length than the target words.[1] In addition, we matched the target words for length (*M* = 4.53, *SD* = 0.72), each being between 4 and 6 letters long. Under these circumstances, according to the OB1-model (Snell, van Leipsig, et al., 2018), both our transposed words (target word 1 and target word 2) should be very likely to be assigned to their syntactically correct positions, thereby maximising the possibility that readers would fail to detect the transposition, and thus, be more likely to judge these sentences as being grammatically correct. The grammatically correct and incorrect post-target words were also matched for length (*M* = 5.90, *SD* = 1.54), each being between 4 and 9 letters long.[2]

According to OB1 Reader model simulations (Snell, van Leipsig, et al., 2018), a word with higher lexical frequency in the parafovea should inhibit processing of the currently fixated word. Additionally, lexical frequency is known to have a robust effect on oculomotor behavior as early as the first fixation during first-pass reading on a word (e.g., Rayner, 2009). Therefore, we controlled for the potential confounding effects of lexical frequency by measuring the Log Zipf lexical frequency (van Heuven et al., 2014) for all words in the sentences and matching the frequencies of the two target words and the pre-target word (see Table 1). Moreover, the post-target words were matched for lexical frequency across the final word grammaticality conditions to ensure any effects on the post-target word were not confounded by a frequency mismatch (see Table 1).[3]

All sentences were pre-screened for naturalness to ensure that the grammatically correct sentences were viewed as natural while the sentences with one or two grammaticality violations were viewed as unnatural (see Table 1). The ratings were provided by 10 native English students or staff from the University of Central Lancashire who had no known reading impairments, had normal or corrected-to-normal vision, and did not take part in the eye-tracking experiment (*M* = 20.9 years, *SD* = 3.6, range = 19–30, Female = 6). The ratings were given for each sentence in each of its four versions from 1 (extremely unlikely to hear this sentence in an everyday conversation) to 7 (extremely likely to hear this sentence in an everyday conversation).

Predictability, similarly, to lexical frequency is known to exert robust effects on reading times measures (e.g., Clifton et al., 2016). Therefore, to ensure that there was no significant difference in predictability between the two target words, 10 additional native English students or staff from the University of Central Lancashire with no known reading impairments and normal or corrected-to-normal vision (*M* = 21.4 years, *SD* = 2.7, range = 19–28, Female = 7) were asked to take part in a close probability task (see

[1] In two sentences out of the 108 sentences, due to human error, the pre-target word had the same length as the two target words. For this reason, we did not include data from trials with these stimuli in our analyses.

[2] Two further sentences had post-target words that differed by one letter between the grammatical and ungrammatical conditions. Thus, the analyses we report here are based on 104 stimuli. Note that with this number of stimuli we still attained the power specifications detailed earlier. For the results based on the full set of 108 stimuli see Supplementary Materials, Tables 1-5.

[3] We note that one aspect of the Post-Target words we did not control for was their dominant syntactic category (as per van Heuven, et al., 2014) such that only seven out of the 104 grammatical post-target words were verbs while 43 of the ungrammatical post-target words were verbs.

Condition	Sentence
Final grammatical word & Non-Transposed targets	Today Paul would prove himself.
Final grammatical word & Transposed targets	Today Paul prove would himself.
Final ungrammatical word & Non-Transposed targets	Today Paul would prove looking.
Final ungrammatical word & Transposed targets	Today Paul prove would looking.

Fig. 1. An example of the stimuli used in Experiment 1. The five regions of interest are indicated as follows: (1) The First Word is presented with a single underline; (2) the Pre-Target Word is presented with a dashed underline; (3) Target Word 1 and Target Word 2 are presented with a double underline; (4) the Post-Target Word is presented with a dotted underline. None of the words and sentences were underlined within the actual experiment.

Table 1
Descriptive statistics for frequency for words at all positions in the sentences, predictability of Target Word 1 and Target Word 2, as well as naturalness ratings for sentences in the four conditions. Frequency data were obtained and calculated via the Log Zipf scale (van Heuven et al., 2014).

Word	Frequency		*t*-test results for the critical comparison between words				
	Range	*M (SD)*	First Word t-value (p-value)	Pre-Target t-value (p-value)	Target Word 1 t-value (p-value)	Target Word 2 t-value (p-value)	Post-Target Grammatical t-value (p-value)
First Word	2.11 – 7.67	5.77 (1.45)					
Pre-Target	2.95 – 6.55	5.32 (0.69)	2.85 (0.004)				
Target Word 1	3.40 – 6.90	5.50 (0.71)	1.72 (0.087)	−1.83 (0.070)			
Target Word 2	3.51 – 7.19	5.48 (0.81)	1.82 (0.071)	−1.47 (0.142)	−0.22 (0.824)		
Post-Target Grammatical	2.37 – 6.45	4.81 (0.80)			6.59 (<.001)	5.98 (<.001)	
Post-Target Ungrammatical	2.39 – 6.35	4.62 (0.74)			8.76 (<.001)	7.98 (<.001)	−1.81 (0.072)
Sentence Type			Naturalness Ratings				
			Range	*M (SD)*			
Grammatically correct sentences			4 – 6.6	5.54 (0.59)			
Sentences with a word transposition and a grammatical final word			1 – 3.1	1.35 (0.43)			
Sentences without a word transposition and an ungrammatical final word			1 – 2.8	1.18 (0.35)			
Sentences with a transposition and an ungrammatical final word			1 – 1.5	1.02 (0.09)			
Target Word			Predictability Ratings				
			Range	*M (SD)*			
Target Word 1			0 – 30	1.7 (5.1)			
Target Word 2			0 – 40	2.6 (7.1)			

Note: *t*-statistics and descriptive values for the final set of 104 sentences that were used for statistical analyses;
Predictability ratings are provided in percentages.

Table 1). Each participant was presented with the first two or three words of each sentence and asked to continue each sentence with the first word that came to their mind. As each participant saw the first word and pre-target twice, 40 filler sentence beginnings were inserted into the questionnaire to mitigate any potential confounds.

Repeated measures *t*-test analyses in R (v4.3.1; R Core Team, 2023) showed that the sentences that contained a word transposition were rated as significantly more natural when the final word was grammatical ($M = 1.35$; $SD = 0.43$) than when the final word was ungrammatical ($M = 1.02$; $SD = 0.09$; $t = 7.54$, $p < 0.001$). Additionally, the sentences without a word transposition were given higher naturalness ratings when the final word was grammatical ($M = 5.54$, $SD = 0.59$) than when the final word was ungrammatical ($M = 1.18$; $SD = 0.35$; $t = 65.50$, $p < 0.001$). Moreover, sentences with a final grammatical word were rated as more natural when they did not contain a transposition than when the two target words were transposed ($t = -58.92$, $p < 0.001$). Similarly, sentences with a final ungrammatical word were rated as more natural when they did not contain a word transposition versus when there was a word transposition as well ($t = -4.41$, $p < 0.001$). It is worth noting that grammatically correct sentences were rated as more natural than sentences with a word transposition only or a final ungrammatical word only ($t = -58.25$, $p < 0.001$ and $t = 64.67$, $p < 0.001$ respectively) which in turn were deemed more natural than sentences containing both a word transposition and a final ungrammatical word ($t = -4.41$, $p < 0.001$ and $t = 7.54$, $p < 0.001$ respectively). These results show a clear gradation in the perceived (un)naturalness between sentences containing no violation, sentences containing one violation (i.e., word transposition or final ungrammatical word), and sentences containing two violations (i.e., word transposition and a final ungrammatical word) to syntax. Importantly, the sentences that had a final grammatical word and did not contain a word transposition (i.e., grammatically correct sentences) were rated as natural while sentences containing one grammaticality violation (i.e., a word transposition or a final ungrammatical word) and sentences containing both a word transposition and a final ungrammatical word were rated as unnatural. These results suggested that the transposition and the ungrammatical final word were clearly seen as a violation.

The repeated measures *t*-test on the predictability norming study data showed that the two targets were unpredictable and that there was no significant difference in their cloze probabilities (target word 1, $M = 1.7$ %; $SD = 5.1$ %; target word 2, $M = 2.6$ %; $SD = 7.1$ %; $t = -1.01$, $p = 0.314$) suggesting that any effect on the targets cannot be attributed to predictability.[4]

4.6. Procedure

Participants were presented with an information sheet and a consent form upon arriving at the laboratory. Next, they were seated in front of a monitor and asked to place their chin and forehead on a headrest for head stabilisation purposes. Once the participants read the instructions, their eyes were calibrated via a 3-point horizontal calibration. This calibration was carried out after each break and whenever necessary, during the experiment.

Each trial started with a drift check 50 pixels (approximately 1 degree of visual angle) to the left of the first letter of the first word in the sentence. In the case of an error above the 0.3-degree threshold, calibration was carried out again. Following the drift check, a fixation cross was presented in the same location, and participants had to look at the cross for 500 ms before the sentence would appear. If a participant moved their eyes and did not look at the cross for 500 ms from the onset of its presentation, a new fixation cross was presented, or a recalibration was conducted. Participants read a block of 12 practice trials followed by 108 experimental trials separated into four blocks of 27 trials. Breaks were provided between each block and whenever the participant required.

Participants were asked to read each sentence to the best of their ability and to determine whether the sentence they just read was grammatically correct or incorrect. In other words, readers had as much time as they needed to reach a decision, and sentences remained on the screen until participants made their response. They used a response box with the left button as the *no* answer and the right button as the *yes* answer. When the participants completed the last trial, they were presented with a *Thank you* message and a debrief form. After receiving their debrief form, participants left the laboratory, and their anonymised data were saved. The entire experimental session lasted on average between 45 and 60 min. The experiment was approved by the Ethics Committee at the School of Psychology and Computer Science at the University of Central Lancashire (Ethics Reference: SCIENCE 0150).

5. Results

5.1. Data Analyses

Data were extracted in Data Viewer v4.1.211 (SR Research), and all analyses were conducted in R v4.3.1 (R Core Team, 2023) with the lme4 package v1.1–33 (Bates et al., 2015). Only fixations between 80 and 800 ms that were not preceded, or followed by, a blink were considered for the analyses. Raw eye movement data were analysed for first fixation duration (FFD: the duration of the first fixation on a word), single fixation duration (SFD: the duration of the fixation when only one fixation was made on a word during first-pass reading), gaze duration (GD: the sum of all fixations on a word during first-pass reading before the eyes move away to another region), and total viewing time (TVT: the total time spent reading the word) for each region within the sentence (i.e., First Word, Pre-Target, Target Word 1, Target Word 2, Post-Target). Go-past time (GPT: the duration of all fixations on the word and any regressive fixations on words to the left of the current word, before the eyes move to the right), skipping (SP: the probability of skipping a word during first-pass reading), and the probability to regress out of a word (Rout: the probability that the eyes move from the currently fixated word to another word to the left during first-pass reading) were analysed for all words within the sentence except First Word. Being the first region of interest within the sentence, no meaningful regressions could be made out of First Word (i.e., gaze duration and go-past time coincided). Additionally, as participants had to make a rightward saccade from the initial fixation cross location to the First Word, parafoveal processing of the First Word was likely different than for other regions of interest in the sentence, and thus the skipping rate would not have been comparable. We analysed fixation probability (FP: the probability to fixate a word) instead of skipping probability for the Post-Target region only since there were no words to the right and thus not fixating the Post-Target would not be comparable to skipping words in sentence-internal regions (i.e., Pre-Target, Target Word 1, or Target Word 2).

We used successive differences contrasts via the *contr.sdif* function in the *MASS* package for R (Venables & Ripley, 2002). With these contrasts we compared each successive level of each factor to the previous level of the same factor (e.g., transposed – not transposed). The *glmer* function with distribution specified as *gamma* and the link specified as *identity* was used to analyse response times and the continuous eye movement variables (first fixation duration, single fixation duration, gaze duration, go-past time, total viewing time). Accuracy and the categorical eye movement variables (skipping probability, fixation probability, regression out probability) were analysed with the *glmer* function, the distribution specified as *binomial*, and the link specified as *logit* (see also Lo & Andrews 2015; Veldre et al., 2017).

Given the assumptions of the SEAM model (Rabe et al., 2024), we might observe the earliest effect of word transposition on the First Word (i.e., N + 2 parafoveal-on-foveal effects) and/or Pre-Target (i.e., N + 1 parafoveal-on-foveal effects). However, all models (SEAM, Rabe et al., 2024; OB1 Reader, Snell, van Leipsig et al., 2018; E-Z Reader, Reichle et al., 2011) agree that we do not expect the grammaticality of the final word (Post-Target) to have an effect on (at least the early) reading times on First Word and Pre-Target.

[4] The results reported here are for the set of 104 stimuli used for the analyses reported in the results section of this Chapter. For the results from the pre-screens based on the full set of 108 stimuli see Supplementary Table 2.

Therefore, only word transposition was included as a fixed factor in the models we carried out for First Word and Pre-Target. We also included word transposition as the only fixed factor in the analysis models we conducted for Target Word 1 and Target Word 2, as this factor was critical for testing the OB1 Reader in comparison with SEAM and E-Z Reader. At these regions, an effect of transposition was predicted by the latter models of reading but not the former. In addition, again to test the predictions of the SEAM model about parallel processing, we also compared grammatical sentences against sentences containing a final ungrammatical word only. If processing proceeds in parallel as per the SEAM assumptions, then the effect of final word grammaticality might show the earliest influence at these regions, on Target Word 1 (N + 2 parafoveal-on-foveal effects) and/or Target Word 2 (N + 1 parafoveal-on-foveal effects).

For the final word (Post-Target), we included in the models both word transposition and final word grammatical fit, as well as their interaction. The word transposition and the interaction factors allowed us to further test the assumption of OB1 Reader versus the other two models of reading we discussed. Any spillover effect of the transposed word pair that we might observe at the Post-Target region might fit with the latter reading models (E-Z Reader and SEAM) but not the former (OB1). The final word grammatical fit factor allowed us to examine the time course of the disruption caused by the ungrammatical final word, and therefore test SEAM assumptions of an effect on not just early, but also later eye movement measures. To test the effect of two, versus one, cue to ungrammaticality and to examine the contribuiting influence of different types of syntactic violation on any observable disruption on the Post-Target, we conducted an additional model with a single three-level factor using a treatment contrast (two violations versus transposition only and two violations versus final word ungrammatical only) as the only fixed factor. Finally, for accuracy and response times we included in the models both word transposition and final word grammatical fit, as well as their interaction to test OB1 Reader assumptions that any

Table 2
Descriptive Statistics for all measures.

Measure/ Condition	Grammatical and Not Transposed	Grammatical and Transposed	Ungrammatical and Not Transposed	Ungrammatical and Transposed
Accuracy	93 (26)	92 (27)	91 (28)	99 (10)
RTs	2148 (779)	2268 (828)	2255 (805)	2159 (784)
	Not Transposed			Transposed
		First Word		
FFD	182 (54)			180 (53)
SFD	179 (54)			177 (52)
GD	202 (93)			206 (102)
TVT	234 (130)			228 (125)
		Pre-Target		
FFD	197 (53)			194 (52)
SFD	204 (53)			201 (52)
GD	240 (88)			238 (89)
SP	8 (26)			8 (28)
ROut	4 (21)			5 (21)
GPT	252 (103)			249 (105)
TVT	338 (188)			334 (197)
	Grammatical and Not Transposed	Grammatical and Transposed	Ungrammatical and Not Transposed	Ungrammatical and Transposed
		Target Word 1		
FFD	219 (63)	223 (69)	218 (60)	222 (69)
SFD	230 (60)	245 (78)	229 (65)	241 (77)
GD	255 (90)	276 (111)	257 (92)	267 (104)
SP	13 (34)	13 (33)	13 (34)	12 (33)
ROut	6 (24)	8 (27)	4 (20)	8 (27)
GPT	274 (117)	303 (145)	269 (116)	296 (145)
TVT	397 (231)	464 (257)	372 (222)	434 (245)
		Target Word 2		
FFD	250 (83)	253 (84)	250 (83)	251 (90)
SFD	282 (95)	300 (98)	287 (95)	293 (99)
GD	328 (144)	345 (171)	332 (148)	328 (155)
SP	8 (27)	7 (25)	9 (29)	7 (25)
ROut	24 (43)	34 (47)	13 (33)	29 (45)
GPT	429 (270)	537 (316)	391 (222)	469 (264)
TVT	466 (236)	544 (267)	518 (255)	504 (255)
		Post-Target		
FFD	243 (95)	243 (105)	257 (95)	237 (95)
SFD	250 (102)	246 (113)	272 (105)	249 (112)
GD	310 (160)	307 (175)	367 (190)	303 (165)
FP	79 (41)	72 (45)	90 (30)	84 (37)
ROut	69 (46)	63 (48)	70 (46)	57 (50)
GPT	763 (573)	790 (641)	848 (600)	679 (555)
TVT	379 (243)	368 (240)	457 (244)	369 (219)

Note. Values for RTs, FFD, SFD, GD, GPT, and TVT are given in milliseconds while values for Accuracy, SP, FP, and ROut are given in percentages. Standard deviations are provided in parentheses.

effect of word transposition and any interaction should only be significant on these measures.

All models were equipped with a full random structure, with both random intercepts and slopes for subjects and items as per Barr et al. (2013). In the event that a model failed to converge, for the models with a single fixed effects factor, we systematically reduced the random structure by first trimming down the level correlation and then removing the slope, starting with the items random structure and then the participants random structure. For the interactive models (on accuracy, response times, and for the Post-Target word), we initially trimmed down the level correlation, and then proceeded to remove interactions and finally slopes, starting with the items random structure first and where needed, continuing with the participants random structure. For the full list of final models that converged, please refer to the OSF page of the project.

For both RTs and eye movement analyses, only observations from correct response trials were considered. Observations more than 2.5 standard deviations away from the mean were removed prior to analyses for all continuous local and global measures. Additionally, all analyses were adjusted via the Bonferroni Correction as per von der Malsburg and Angele (2017) to mitigate for type 1 statistical errors. Since the final number of observations varied across analyses, further information regarding the number of observations used for each model is provided within the R code, available on the OSF.

5.2. Accuracy results

Overall, in respect of accuracy, participants performed the task well gaining accuracy scores of more than 90 % in all conditions (Table 2). Despite the good overall performance on the task, there were robust differences across conditions. There was an effect of transposition on accuracy ($\beta = 1.40$; $z = 3.20$), and an effect of whether the final word was, or was not, grammatical ($\beta = -1.70$; $z = -3.42$). As can be seen from the means in Table 2, these effects were largely driven by the interactive pattern ($\beta = -3.15$; $z = -3.43$) such that grammaticality decision performance was substantially higher when sentences contained both a transposition and an ungrammatical final word than any of the other three conditions for which performance was very similar.

We note that in considering these results, we are comparing conditions in which participants were required to make *yes* (i.e., grammatical) decisions with *no* (i.e., ungrammatical) decisions, and ordinarily such comparisons across decision types are not made (e.g., in lexical decision tasks). However, here we feel that such comparisons are important and necessary because they allow us to evaluate whether there are performance differences in the context of more, or less, evidence in favor of an outcome. That is to say, in making these comparisons, we can assess whether having three, two, one, or no cues to grammaticality violations within a sentence influences decision accuracy. And, as can be seen, it is clearly the case that having three grammatically illegal words in a sentence causes participants to make more accurate judgments than having two or only one ungrammatical word in a sentence (see also Spinelli et al., 2024). This result is interesting in that it suggests that evidence to make a *no* (i.e., ungrammatical) decision is cumulative such that when there are two cues to an ungrammatical sentence, that is, a transposition and an ungrammatical final word, then participants are better at forming an appropriate judgment than when only one of the two cues is present. At a more general level, the accuracy results suggest that transpositions may be easier to detect than was previously shown by Mirault et al. (2018; 2020), and hypothesised by the OB1 Reader (Snell, van Leipzig, et al., 2018). There might be at least three reasons why responses were more accurate for ungrammatical sentences and less accurate for grammatical in the present study than was the case in the Mirault et al. (2018; 2020) and Spinelli et al. (2024) studies. First, in Mirault et al., participants were required to make speeded grammaticality decisions,[5] whereas in the present study there was no time limit in respect of decisions (accordingly, the summed total times for each region and the response times in the current experiment indicate that total times were longer here than was the case in the Mirault et al. studies). Perhaps the requirement to make more rapid responses in the Mirault et al. and Spinelli et al. studies may have led participants to make more errors as per standard speed-accuracy relations. However, as observed by Liu et al. (2021) and in the current study, readers were still significantly better at judging sentences containing two grammaticality violations as ungrammatical than sentences containing a word transposition only. In other words, the *Transposed-Word* effect reported in previous studies (e.g., Mirault et al., 2018) was still observed in the present study despite the fact that grammaticality decisions were unspeeded. Second, due to our fully rotated experimental design, 75 % of our stimuli required a *no* (ungrammatical) decision, whilst 25 % required a *yes* (grammatical decision), thus, some of this effect may arise as a consequence of response bias. In order to further examine whether the accuracy results were influenced by a response bias, we conducted a post-hoc analysis in which we included an additive covariate to the original model that indicated whether each observation belonged to the first or the second half of the experiment to assess whether participants potentially became better at detecting ungrammaticalities as they progressed through the experiment. The new model failed to show any effect associated with this covariate ($\beta = 0.16$; $z = 1.10$), suggesting that response bias likely was not the primary driver for our findings on accuracy. Third, the transpositions in the current study were always formed from two quite long words (4–6 letters) and most often at least one of these words was a content word. In contrast, transposed word pairs in Mirault et al. (2018; 2020) (and in Huang & Staub, 2021a) were often formed from at least one shorter function word. It is possible that sensitivity to transpositions, and therefore accuracy of response, may have been reduced in the latter relative to the former situation (see Wen et al., 2021).

[5] We note that although Mirault et al. (2018, 2020) asked participants to make a speeded response, there was no time limit for participants to respond, and the sentence remained on the screen until a response was made.Spinelli et al. (2024. Experiment 1) did include a time out, but this was set at 20 s to approximate the lack of a time limit in the studies by Mirault et al. (2018, 2020).

5.3. Response time results

Overall, participants took on average just over 2 s to form their grammaticality decisions to the sentences ($M = 2206$ ms; $SD = 800$ ms; *Range* = 863–7095 ms; see Table 2). The effects of transposition and of final word grammaticality did not reach significance likely because of the very robust crossover interaction ($\beta = 248.06$; $z = 27.77$), such that for sentences with a transposition, response times were longer when they contained a final grammatical than ungrammatical word (as in Mirault et al., 2018, 2020 and Spinelli et al., 2024), while for sentences without a transposition, response times were shorter when they were grammatical than ungrammatical (as in Spinelli et al., 2024). Hence, it is clear that participants were fastest in reaching a grammaticality decision for grammatically correct sentences, slightly slower for sentences containing two grammaticality violations, and they were substantially, and comparably slower when reaching a decision for sentences containing a single grammaticality violation. The finding that response times were comparable for sentences containing a word transposition only and sentences containing a final ungrammatical word only is in line with the present findings for accuracy as well as the findings on response times by Spinelli et al. (2024: Experiment 1), and seems to further suggest that the *Transposed-Word* effect reported previously (e.g., Mirault et al., 2018) is likely not driven by flexible word position coding as per OB1 Reader (Snell, van Leipsig, et al., 2018). We, again, note that these response times are somewhat longer than those reported in studies employing speeded grammaticality decision tasks.

5.4. Eye movement measures results

The descriptive statistics for the eye movement measures are shown in Table 2 and GLMMs are shown in Table 3 and Table 4. On the First Word and Pre-Target word, we did not observe any significant effect of word transposition, indicating no evidence of a parafoveal sensitivity to the word order syntactic violation. Therefore on First Word and Pre-Target we found no support for the predictions derived from the SEAM model.

Recall that our manipulations in this experiment did not involve changes to the first two words of our sentences. That is to say, up until Target Word 1, the sentence was completely grammatical and made perfect sense in all conditions. However, Target Word 1 was the first word in the sentence that was (in some conditions) ungrammatical, and unsurprisingly, patterns of effects were markedly different at this word, where we obtained very early, substantive effects in our reading time measures. For Target Word 1, there was a significant effect of transposition on single fixation duration ($\beta = 12.59$; $z = 2.56$), gaze duration ($\beta = 17.64$; $z = 3.66$), go-past time ($\beta = 31.98$; $z = 5.92$), and total viewing time ($\beta = 72.47$; $z = 10.55$), with longer reading times when the two target words were transposed, compared to when they were not. There was also an effect of transposition on regression out probability ($\beta = 0.48$; $z = 3.50$) such that participants were more likely to regress out of Target Word 1 when the two targets were transposed compared with when they were not.

All of these results indicate that the transposition of the two target words produced rapid and very substantive disruption to processing upon fixation of Target Word 1. Note also that this was the case in the context of participants correctly identifying that sentences containing a transposition were ungrammatical on over 90 % of trials. Clearly, in the present study, participants were quite able to detect transpositions and to judge that sentences were ungrammatical when they fixated the first word of the transposed word pair. Furthermore, transpositions clearly had a rapid influence on eye movement behavior. These results are entirely in line with serial processing accounts such as the E-Z Reader model (Reichle et al, 1998; Reichle, 2011). These results might also be explained by parallel models such as SEAM (Rabe et al., 2024) in that a transposed word pair resulting in an ungrammatical sentence should produce disruption to reading when it is processed. However, SEAM might have difficulty explaining an increased probability of regressing out and inflated go-past time for the transposed compared to non-tranposed Target Word 1 since the model assumes that regressions out are more likely, and go-past time longer, for a grammatical than an ungrammatical sequence. Finally, these results are perhaps most problematic for the OB1 model (Snell, van Leipsig, et al., 2018) that specifies that words are identified in parallel and integrated together at a later stage of processing based on a spatiotopic map. If this was the case, then it is unclear why participants detected word transpositions on between 92–99 % of trials, nor is it clear why transpositions caused substantive and immediate disruption to eye movements upon fixation of the first word of the transposed word pair.

With respect to the comparison between grammatically correct sentences and sentences containing only a final ungrammatical word, total viewing time approached significance ($\beta = 27.40$; $z = 2.32$). Participants spent longer re-reading Target Word 1 when the final word was grammatical than ungrammatical. This pattern of late effects presumably reflects increased checking for grammatical sentences or an increased likelihood to launch regressions for grammatical than ungrammatical (Post-Target) words as per SEAM, although we note that these results should be taken with caution as they only approached significance.

On Target Word 2 there was a significant effect of transposition on go-past time ($\beta = 93.61$; $z = 16.57$), and total viewing time ($\beta = 36.54$; $z = 5.85$), an effect that approached significance on single fixation duration ($\beta = 15.03$; $z = 2.22$). Reading times were longer when the two targets were transposed, compared to not transposed. This pattern of effects is very similar to that observed for Target Word 1 and suggests disruption to eye movements due to the word transposition and resulting ungrammaticality. There was also an effect of transposition on regression out probability ($\beta = 0.82$; $z = 6.99$), such that participants were more likely to regress out of Target Word 2 when the two targets were transposed compared with when they were not transposed.

Additionally, on Target Word 2, there was a significant effect when comparing grammatical sentences and sentences with a final ungrammatical word for go-past times ($\beta = 34.75$; $z = 2.63$), regression out probability ($\beta = 0.87$; $z = 5.37$), and total viewing time ($\beta = -50.97$; $z = -3.70$). Go-past times were shorter, probability of regressions out lower, and total viewing times longer for sentences with a final ungrammatical word and no transposition compared to grammatical sentences. This pattern of results indicates that the (un) grammaticality of the final word was detected from the parafovea and influenced processing on the Target Word 2 region, and this

Table 3
Fixed effects estimates from the generalised mixed effects models.

Measure/ Condition	Intercept				Transposition (Transposed-Not Transposed)				Final Word Grammaticality (Grammatical-Ungrammatical)				Interaction			
	β	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*
Accuracy	4.04	0.24	**16.86**	[3.57, 4.51]	1.40	0.44	**3.20**	[0.54, 2.26]	−1.70	0.50	**−3.42**	[-2.67, −0.72]	−3.15	0.92	**−3.43**	[-4.95, −1.35]
RTs	2410.50	12.50	**192.90**	[2386.01,2434.89]	35.31	15.07	2.34	[5.77,64.84]	15.79	7.16	2.21	[1.76,29.82]	248.06	8.93	**27.77**	[230.55,265.57]
First Word																
	β	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z) −value*	*Conf Int*	*β*	*SE*	*t(z) −value*	*Conf Int*	*β*	*SE*	*t(z) −value*	*Conf Int*
FFD	182.58	6.49	**28.13**	[169.85,195.30]	−2.10	1.85	−1.14	[-5.72,1.53]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
SFD	192.81	10.02	**19.24**	[173.17,212.45]	−2.85	2.33	−1.22	[-7.43,1.72]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
GD	207.14	7.37	**28.11**	[192.70,221.59]	2.82	2.92	0.97	[-2.91,8.55]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
TVT	234.42	7.35	**31.91**	[220.02, 248.82]	−7.60	4.76	−1.60	[-16.92,1.72]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Pre-Target																
	β	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*
FFD	200.18	5.50	**36.42**	[189.41,210.96]	−3.49	2.45	−1.43	[-8.28,1.30]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
SFD	210.49	6.33	**33.25**	[198.09,222.90]	−2.72	3.42	−0.80	[-9.41,3.97]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
GD	242.67	8.01	**30.30**	[226.97,258.36]	−2.50	3.88	−0.65	[-10.10,5.09]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
SP	−3.36	0.26	**−13.17**	[-3.86, −2.86]	−0.27	0.20	−1.35	[-0.66,0.12]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
ROut	−3.60	0.19	**−18.59**	[-3.99, −3.22]	0.04	0.17	0.25	[-0.28,0.37]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
GPT	257.69	6.53	**39.48**	[244.9,270.49]	−1.61	4.48	−0.36	[-10.38,7.17]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
TVT	346.37	9.17	**37.76**	[328.39,364.35]	0.61	7.73	0.08	[-14.54,15.68]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Target Word 1																
	β	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*
FFD	223.46	5.30	**42.16**	[213.07,233.84]	4.84	3.65	1.33	[-2.32,12.00]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
SFD	238.98	7.62	**31.37**	[224.05,253.92]	12.59	4.92	**2.56**	[2.94,22.23]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
GD	263.99	5.47	**48.25**	[253.27,274,71]	17.64	4.82	**3.66**	[8.20,27.09]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
SP	−2.60	0.23	**−11.40**	[-3.04, −2.15]	−0.05	0.11	−0.52	[-0.26,0.15]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
ROut	−2.94	0.14	**−20.53**	[-3.22, −2.66]	0.48	0.14	**3.50**	[0.21,0.74]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
GPT	288.16	6.46	**44.64**	[275.50,300.81]	31.98	5.40	**5.92**	[21.39,42.57]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
TVT	425.77	11.57	**36.79**	[403.09,448.45]	72.47	6.87	**10.55**	[59.00,85.94]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Target Word 2																
	β	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*
FFD	253.92	5.05	**50.33**	[244.03,263.81]	2.02	4.63	0.44	[-7.07,11.10]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
SFD	286.75	8.64	**33.19**	[269.81,303.68]	15.03	6.78	2.22	[1.74,28.32]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
GD	337.08	5.82	**57.92**	[325.68,348.49]	10.27	5.30	1.94	[-0.11,20.65]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
SP	−3.02	0.20	**−15.35**	[-3.41, −2.64]	−0.24	0.17	−1.40	[-0.58,0.10]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
ROut	−1.32	0.13	**−10.13**	[-1.58, −1.06]	0.82	0.12	**6.99**	[0.59,1.05]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
GPT	462.06	8.46	**54.62**	[445.48,478.63]	93.61	5.65	**16.57**	[82.53,104.68]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

(*continued on next page*)

Table 3 (*continued*)

Measure/ Condition	Intercept				Transposition (Transposed-Not Transposed)				Final Word Grammaticality (Grammatical-Ungrammatical)				Interaction			
	β	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*
TVT	515.55	6.49	**79.44**	[502.83,528.27]	36.54	6.25	**5.85**	[24.30,48.78]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
								Post-Target								
	β	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*
FFD	244.66	7.57	**32.32**	[229.83,259.5]	−13.65	5.59	−2.44	[-24.61, −2.70]	−7.24	5.51	−1.32	[-18.04,3.56]	20.92	10.04	2.08	[1.25,40.59]
SFD	269.79	8.75	**30.84**	[252.65,286.93]	−21.06	8.87	−2.38	[-38.44, −3.68]	−20.13	7.31	**−2.76**	[-34.45,-5.81]	20.78	9.07	2.29	[3.01,38.54]
GD	318.79	6.72	**47.47**	[305.63,331.95]	−39.74	5.70	**−6.97**	[-50.91, −28.56]	−37.36	4.83	**−7.74**	[-46.82,-27.9]	67.15	8.01	**8.39**	[51.45,82.84]
FP	2.04	0.20	**10.11**	[1.65,2.44]	−0.47	0.13	**−3.64**	[-0.72, −0.22]	−0.98	0.13	**−7.34**	[-1.24,-0.72]	0.30	0.20	1.54	[-0.08,0.69]
ROut	0.73	0.17	**4.34**	[0.40,1.07]	−0.53	0.08	**−6.28**	[-0.69, −0.36]	0.09	0.12	0.77	[-0.14,0.32]	0.42	0.17	**2.54**	[0.10,0.75]
GPT	756.58	6.87	**110.11**	[743.12,770.05]	−82.65	7.68	**−10.76**	[-97.70, −67.59]	−16.93	6.22	**−2.72**	[-29.12,-4.75]	197.21	12.82	**15.38**	[172.08,222.33]
TVT	388.87	8.05	**48.29**	[373.09,404.66]	−57.95	6.21	**−9.34**	[-70.12, −45.79]	−56.37	6.23	**−9.05**	[-68.58,-44.16]	85.72	7.65	**11.21**	[70.74,100.71]

Note. Significant terms are presented in bold, and terms approaching significance are underlined. All *p* values were adjusted via the Bonferroni correction with the summary(glht(<modelName >), test = adjusted(“bonferoni”)) function. N/A signifies that the factor or interaction was not considered in the model.
ˆConfidence intervals (Conf Int) are calculated with the *confint* function in R with method = *Wald* at the 2.5 percentile and 97.5 percentile.

Table 4

Fixed effects estimates from the specific pairwise comparisons on Target Word 1; Target Word 2 and Post-Target Word.

Measure/ Condition	Intercept				Grammatically Correct versus Final Ungrammatical Word Only				Transposed Word Pair and Final Ungrammatical Word versus Transposed Word Pair Only				Transposed Word Pair and Final Ungrammatical Word versus Final Ungrammatical Word Only			
								Target Word 1								
	β	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*
FFD	219.00	6.37	**34.38**	[206.52,231.49]	0.41	4.36	0.10	[-8.13,8.95]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
SFD	230.82	7.07	**32.65**	[216.96,244.67]	−4.17	6.17	−0.68	[-16.25,7.92]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
GD	252.94	7.08	**35.73**	[239.07,266.82]	−1.73	5.89	−0.29	[-13.27,9.82]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
SP	−2.56	0.24	**−10.75**	[-3.02, −2.09]	−0.01	0.15	−0.05	[-0.29,0.28]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
GPT	271.27	6.64	**40.89**	[258.27,284.27]	7.97	7.23	1.10	[-6.20,22.15]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
TVT	384.98	15.96	**24.11**	[353.69,416.27]	27.40	11.80	2.32	[4.27,50.53]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
								Target Word 2								
	β	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*
FFD	252.45	7.02	**35.97**	[238.70,266.21]	0.63	5.52	0.12	[-10.18,11.44]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
SFD	289.13	10.42	**27.74**	[268.71,309.56]	−6.08	13.84	−0.44	[−33.20,21.04]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
GD	332.21	6.57	**50.57**	[319.33,345.08]	−2.60	7.30	−0.36	[-16.90,11.71]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
SP	−3.07	0.25	**−12.24**	[-3.57, −2.58]	−0.04	0.27	−0.16	[-0.57,0.48]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
ROut	−1.79	0.16	**−11.49**	[-2.10, −1.49]	0.87	0.16	**5.37**	[0.55,1.19]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
GPT	403.37	11.28	**35.75**	[381.25,425.48]	34.75	13.24	**2.63**	[8.80,60.71]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
TVT	498.45	10.91	**45.70**	[477.07,519.83]	−50.97	13.78	**−3.70**	[-77.99, −23.96]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
								Post-Target								
	β	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*
FFD	237.05	6.25	**37.93**	[224.80,249.29]	N/A	N/A	N/A	N/A	1.93	7.37	0.26	[-12.50,16.37]	22.94	6.41	**3.58**	[10.38,35.50]
SFD	261.77	10.54	**24.84**	[241.11,282.43]	N/A	N/A	N/A	N/A	−11.13	8.37	−1.33	[-27.54,5.28]	34.74	8.43	**4.12**	[18.23,51.26]
GD	300.03	8.67	**34.61**	[283.04,317.02]	N/A	N/A	N/A	N/A	0.75	6.62	0.11	[-12.23,13.73]	69.30	8.00	**8.66**	[53.61,84.98]
FP	2.17	0.22	**9.83**	[1.74,2.61]	N/A	N/A	N/A	N/A	−0.86	0.12	**−6.94**	[-1.10,-0.62]	0.72	0.15	**4.84**	[0.43,1.01]
ROut	0.34	0.19	1.74	[-0.04,0.72]	N/A	N/A	N/A	N/A	0.32	0.12	**2.71**	[0.09,0.55]	0.73	0.11	**6.42**	[0.51,0.95]
GPT	672.09	11.81	**56.90**	[648.94,695.24]	N/A	N/A	N/A	N/A	76.55	10.77	**7.11**	[55.44,97.65]	175.32	9.83	**17.84**	[156.06,194.58]
TVT	362.06	11.69	**30.99**	[339.16,384.96]	N/A	N/A	N/A	N/A	−8.61	9.75	−0.88	[-27.72,10.51]	103.04	9.55	**10.79**	[84.32,121.76]

Note. Significant terms are presented in bold, and terms approaching significance are underlined. All *p* values were adjusted via the Bonferroni correction with the summary(glht(<modelName >), test = adjusted(“bonferoni”)) function. N/A signifies that the factor or interaction was not considered in the model.

ˆConfidence intervals (Conf Int) are calculated with the *confint* function in R with method = *Wald* at the 2.5 percentile and 97.5 percentile.

[1]We do not report model statistics for the probability to regress out of Target Word 1 when comparing grammatically correct sentences versus sentences containing a final ungrammatical word only because the model failed to converge even with the minimal random structure for items and participants.

caused readers to move their eyes forward in order to fixate the final word of the sentence and likely more often return to Target Word 2 when the Post-Target was ungrammatical. The go-past time and regression out probability results might fit with SEAM assumptions but are inconsistent with the serial processing assumptions of E-Z Reader (Reichle et al., 1998; Reichle, 2011) as well as the parallel processing assumptions of OB1 Reader (Snell, van Leipzig, et al., 2018). All models however might be able to explain the results for the total viewing time as reflecting disruption during later processing associated with a syntactic violation.

Let us next consider the results for the Post-Target word. Here the effect of transposition was significant for gaze duration (β = -39.74; z = -6.97), go-past time (β = -82.65; z = -10.76), total viewing time (β = -57.95; z = -9.34), and approached significance for first fixation duration (β = -13.65; z = -2.44) and single fixation duration (β = -21.06; z = -2.38) with longer reading times when the two targets were not transposed versus transposed. this effect was also significant for fixation probability (β = -0.47; z = -3.64) and regression out probability (β = -0.53; z = -6.28), with a lower probability to fixate and to regress out of the Post-Target when the two targets were transposed versus not transposed. We also found an effect of final word grammaticality on single fixation duration (β = -20.13; z = -2.76), gaze duration (β = -37.36; z = -7.74), go-past time (β = -16.93; z = -2.72) and total viewing time (β = -56.37; z = -9.05), with longer reading times when the final word was ungrammatical than grammatical, as well as on fixation probability (β = -0.98; z = -7.34) with a higher probability to fixate the sentence final word when it was ungrammatical than grammatical.

The effects of word transposition and final word grammaticality were qualified by interactions that were robust for gaze duration (β = 67.15; z = 8.39), go-past time (β = 197.21; z = 15.38), total viewing time (β = 85.72; z = 11.21), and regression out probability (β = 0.42; z = 2.54), approached significance for single fixation duration (β = 20.78; z = 2.29), and patterned numerically similarly for first fixation duration (β = 20.92; z = 2.08), and fixation probability (β = 0.30; z = 1.54). In the reading time measures, these interactions were largely driven by longer times for the Post-Target word when it was ungrammatical and the sentences did not contain a transposition. Note that this was particularly the case for go-past time. Readers spent longest re-reading sentences with an ungrammatical final word but no transposition (and note also that this effect complements the total viewing time effect observed for the Target Word 2 region). These results are consistent with participants experiencing disruption to processing when they encountered an ungrammatical word for the first time. In the three other conditions, readers had either already encountered an ungrammaticality (i.e., a transposed word pair) earlier in the sentence, and if so had already detected the anomaly and reacted to it (by taking longer to read the word, regressing and re-reading early portions of the sentence), or alternatively, they had processed a sentence that was perfectly grammatical up to the final word and had, therefore, not experienced difficulty until this point in the sentence. Two other differences seem to contribute significantly to the interactive effects, namely, the relatively low rate of regressions from the final word and the reduced go-past times when the sentence contained both a transposed word pair and an ungrammatical final word relative to all other conditions. As suggested previously, it seems very likely that values for these measures were reduced because participants had already detected the ungrammaticality, and probably experienced disruption to processing, as a result of encountering the transposed word pair earlier in the sentence. Alternatively, this reduction in go-past time and decrease in probability to regress out of the Post-Target when this word was preceded by a word transposition may be explained by SEAM. According to this model, an ungrammatical word cannot be integrated into context, resulting in fewer regressions out compared to a grammatical word. However, if this was the case, then it it puzzling that we do not observe a similar pattern for those sentences with a final ungrammatical word only compared to grammatical sentences. On the contrary, probability of regressing out of the Post-Target word for these sentences, and the associated go-past time were comparable and longer respectively than for grammatical sentences, a finding that is inconsistent with SEAM assumptions. A more likely possibility is that at the Post-Target word, readers had almost certainly already detected the ungrammaticality associated with the transposition, and consequently, when the Post-Target word provided an additional indication that the sentence was ungrammatical, this had a far reduced disruptive influence on processing (in that it provided a redundant cue to an ungrammatical sentence). In fact, given that overall grammaticality decision times were shortest for the grammatically legal sentences and sentences containing a transposition as well as a grammatically illegal final word, it might even be the case that having two cues to ungrammaticality actually reduced disruption by facilitating readers in forming their decision.

The additional statistical model we ran to test the influence of different types of syntactic violation showed that eye movements for sentences containing two violations (word transposition and final ungrammatical word) were significantly reduced than for sentences containing an ungrammatical final word only. Also, these analyses showed that for most measures, sentences containing two violations produced results that were comparable to those for sentences containing a word transposition only. Specifically, first fixation duration (β = 22.94; z = 3.58), single fixation duration (β = 34.74; z = 4.12), gaze duration (β = 69.30; z = 8.66), go-past times (β = 175.32; z = 17.84), and total viewing times (β = 103.04; z = 10.79) were shorter, and the probability to fixate (β = 0.72; z = 4.84), and to regress out (β = 0.73; z = 6.42) of the Post-Target word were lower, for sentences containing two violations compared to a final ungrammatical word only. Conversely, only go-past times (β = 76.55; z = 7.11) and probability to regress out (β = 0.32; z = 2.71) were decreased, and fixation probability (β = -0.86; z = -6.94) was increased on the Post-Target word for sentences containing two violations compared to sentences containing a word transposition only. Taken together these findings provide further support for our suggestion that an ungrammaticality on the final word likely captured attention at the point of processing Target Word 2 specifically when the two target words were not transposed.

Overall, at a general level, the results from the current experiment do not fit perfectly well with any of the existing computational models of oculomotor control during reading. Perhaps the model that is supported least by the results is the OB1 Reader since we obtained very robust effects of word transposition and these appeared for early measures of processing at Target Word 1. Whilst we obtained a transposed word effects in response times and accuracy for the grammaticality decision task, according to the OB1 model, transposed words should not cause disruption to the eye movement record. This clearly was not the case. In relation to the SEAM model, the regression out probability and go-past times were increased for Target Word 2 when the sentences were grammatically correct relative to when only the final word was ungrammatical. This result may offer support for the SEAM model. However, at the

Post-Target word where SEAM may have predicted comparable effects for sentences with a final ungrammatical word only and those with a transposition and a final ungrammatical word, the patterns of effects (for regression out and go-past times) actually differed. Furthermore, effects of this type that we might have observed should have been reduced relative to those obtained for grammatically correct sentences. However, in fact, the go-past times and regression out probabilities at the Post-Target word for grammatically correct sentences were actually lower than for sentences with a final ungrammatical word only. Finally, while the pattern of effects in later measures for Target Word 2 does not necessarily fit with the E-Z Reader model, most of the other effects observed across the five regions of interest do seem to be consistent with the serial processing perspective.

6. Discussion

In the present experiment we investigated how transposed word pairs and sentence-final words that were ungrammatical affected eye movement behavior as participants processed sentences in order to make a grammaticality decision. Our motivation for the experiment was to examine these two influences through their orthogonal manipulation and to use a set of stimuli that had been well normed to ensure that both manipulations resulted in sentences that participants perceived to be ungrammatical. Particularly, we sought to understand how frequently, and at what point during processing, participants first detected ungrammaticalities and whether, and how, both cues to ungrammaticality affected eye movements as the sentences were read. In doing this, we also evaluated which of the current models of eye movement behavior best accounted for our findings.

In some respects, the results from our study were very clear and straightforward, whilst in other respects, it is probably fair to say that the patterns of results were complex. Before we consider each of our findings in turn, it is perhaps worth commenting on some general characteristics of our experiment that likely contributed to complexities in our findings. The current experiment was based on a series of original experiments that first showed the *Transposed-Word* effect (Mirault et al., 2018; 2020). These original experiments used linguistic stimuli that took the form of 5-word sentences. The third and fourth words of the sentence were potentially transposed, while the final, fifth word of the sentence could either be grammatical or ungrammatical but only when preceded by a transposed word pair. Because we wished to use stimuli that were directly comparable to the types of stimuli that were used in the original experiments, we adopted sentences of this form here. However, it is important to note that there are a number of reasons why stimuli with this type of construction might not be the most ideal to use in an eye movement experiment that investigates the time course of ungrammaticality detection. First, the sentences are short, which means that readers will not make too many eye movements as they read those sentences, and of course, this is not ideal because eye movement behavior is the dependent measure in this experiment. More important, however, is the fact that the two manipulations of ungrammaticality in the sentence, namely, the word pair transposition and the grammaticality of the sentence-final word take place immediately adjacent to each other in the sentence. Given that sentential ungrammaticality produces significant and extended disruption to eye movement behavior during reading, effects associated with the first ungrammaticality manipulation (the transposed word pair) will almost certainly spillover and affect fixations made on the following word, that is, the sentence-final word which itself was manipulated for grammaticality. Often in eye movement experiments that seek to investigate the influence of two different variables on processing, experimenters design their stimuli to ensure that several words appear in a sentence or a text between the portions of the sentence that will likely produce disruption to processing. This is done to make it possible to tease apart and separately observe particular effects in the eye movement record that are associated specifically with each variable. Thus, in the present experiment, the fact that the sentences were only 5 words long in total and that the regions of text associated with the two manipulations of sentence ungrammaticality themselves comprised three of the five words in the sentence, and that these two regions were adjacent to each other, meant that there was a significant degree of congestion with respect to the influences of the variables that we manipulated. Furthermore, ordinarily in eye movement studies investigating reading, experimenters choose not to implement critical experimental manipulations on sentence final words since for such words readers do not have an opportunity to move their eyes to a word to the right and thereby face a decision of regressing or remaining fixated on the final word in the sentence. Furthermore, fixations on the final word of a sentence are associated with sentence wrap-up effects (e.g., Warren et al., 2009) and such effects might likely cloud any experimental effects associated with manipulations of the final word. It is very likely, therefore, that the effects we obtained in the present experiment may not have been as clear as might have been preferred had the sentences been longer and the two manipulations had been spaced apart to a greater degree within the sentence (e.g., 12-word sentences with the transposed word pair appearing at positions 4 and 5 and the ungrammatical word appearing at position 10 within the sentence). Stimuli with such structure may have produced less congested effects and would have allowed for more effective examination of the time course of the two sources of influence on processing. And to reiterate, the reason that we did not adopt sentences that were more optimally formed in these regards was because we wished to use stimuli that allowed direct comparison with previous experiments.

Regardless of the particular form of the sentences we used here, it was certainly the case that our experimental manipulations produced very robust effects, and this was at least in part due to the pre-screening procedures that we adopted prior to conducting the eye movement experiment. Recall, unlike almost all previous studies investigating *Transposed-Word* effects in reading, we assessed the degree to which subjects judged our sentences as natural when our target word pairs were, or were not, transposed and when the sentence final word was, or was not, grammatically illegal. And in line with our expectations, readers rated sentences as unnatural when they contained a transposition, or an ungrammatical final word, and particularly when the sentences contained both. In this way, we could be sure that our transposed word pair and sentence final word manipulations were such that they caused readers to assess the stimuli as ungrammatical as required.

As noted, it was the case that aspects of our results were very clear. To briefly summarize, participants had high levels of accuracy in this experiment, making correct grammaticality decisions on over 90 % of occasions in all conditions, and performing almost perfectly when sentences contained two cues indicating that they were ungrammatical. These grammaticality decision accuracy rates are substantially higher than those that have been previously reported; for example, for the sentences containing only a word transposition, accuracy was between 84 % (in a laboratory experiment) and 88 % (in an on-line experiment) in Mirault et al. (2018), 84 % in Mirault et al. (2020), and between 76 % (when the transposed word pair contained two short open-class words) and 91 % (when the transposed word pair contained one open-class and one closed-class word and one word was short while the other was long) in Huang and Staub (2021a). The increased accuracy rates in the current study very likely reflect the fact that the stimuli in the present experiment were screened to ensure that the word transposition and sentence-final word manipulations caused participants to assess the sentence as ungrammatical. Whether these manipulations were quite so effective in (at least some of the) previous studies is open to question. A second reason for the reduced error rates in the present study is that our experimental task required participants to read the sentences carefully and to understand them to the best of their ability. Previous studies have often required participants to make speeded grammaticality decisions which may have contributed to participants making more errors (see Liu et al., 2021). That said, similarly to Liu and colleagues (2021), we did still find a *Transposed-Word* effect on accuracy such that participants were better at correctly judging sentences with two syntactic violations as ungrammatical (M = 99 %) versus sentences containing a word transposition only (M = 92 %) when grammaticality decisions were not speeded. These findings further support the notion that *Transposed-Word* effects can be observed in an unspeeded grammaticality judgment task and are consistent with existing literature.

Considering the overall response time results, we found these to be shorter when sentences were grammatical than when sentences were ungrammatical. However, when participants decided that sentences were ungrammatical, their decisions were slower when sentences contained one cue to their ungrammaticality (either a transposed word pair, or an ungrammatical final word) than when they contained both cues (both a transposed word pair and an ungrammatical final word). This is consistent with Mirault et al. (2018; 2020), and Spinelli et al. (2024: Experiment 1), as they also found response times to be shorter for sentences containing both a transposition and a final ungrammatical word than sentences containing only the word transposition. Note, though, that in the current study response times were somewhat longer (approximately 2.2–2.3 s) than has been reported in previous studies (approximately 1.3–1.8 s, Mirault et al., 2018; 1.7 s, Mirault et al., 2020). The patterns observed for RTs fit with those observed for accuracy and the *Transposed-Word* effects previously reported in the literature (e.g., Mirault et al., 2018, 2020).

Next, we will consider the eye movement and reading time data for each region of the sentences. When participants started reading the sentences, they did so without any difficulties over the first two words and this was the case for sentences in all conditions. However, as soon as participants fixated the first word of a transposed word pair, they very rapidly experienced disruption to processing relative to when the word pair was not transposed. Thus, at the third word of the 5-word sentence (i.e., Target Word 1 region), there was disruption to first pass reading caused by the third and fourth word having their positions switched. Note also that there was no influence of the grammaticality of the final word of the sentence at this point. These findings are quite consistent with the results reported by Huang and Staub (2021a). Huang and Staub found that when readers did identify that sentences containing a transposed word pair were ungrammatical, there was rapid disruption to processing in the eye movement record, in accord with the current results. These aspects of the current results are straightforward and very clear. From our grammaticality decision data, we know that when our stimuli contained a transposed word pair, participants judged them to be ungrammatical. The eye movement results, therefore, fit perfectly with a serial processing account (the E-Z Reader model) in that there was no influence of the transposed word pair prior to the fixation of the first word of the pair, nor was there a parafoveal influence of the grammaticality of the final word in the sentence at the Target Word 1 region. Whilst the OB1 Reader model might predict a lack of disruption to processing at the first and second word of the sentence, it is not immediately obvious that it can explain the results at Target Word 1 when the sentence contained a transposition. The OB1 Reader model specifies that any ungrammaticality arising due to words appearing out of order should not produce immediate disruption to reading on any local eye movement measures. Recall that OB1 Reader stipulates that words are encoded in relation to a spatiotopic map and through this mechanism, readers are able to process words when they do not appear in their correct grammatical order. Taking grammaticality decisions and reading time effects together, there should be two consequences of this specification that did not occur in response times, nor in relation to the first pass eye movement measures for the Target Word 1 (and Target Word 2) region of the sentence. First, readers should have failed to detect ungrammaticalities arising due to words appearing out of order, and second, readers should not have exhibited disruption to processing in the eye movement record when they encountered transposed words. The current results show that readers almost always detected ungrammaticalities due to word transpositions with accuracy rates of over 90 %. Furthermore, readers showed rapid disruption to processing when they fixated the first word of a transposed word pair. To our minds, therefore, the first-pass reading time results for Target Word 1 of the sentences are inconsistent with the account offered by the OB1 Reader model. Finally, SEAM suggests that words are identified in parallel during reading and stipulates that regressions are triggered when a word needs to be bound together with an already indentified word to form a valid syntactic structure for the sentence. We also note that when readers process a word in a sentence that is ungrammatical, it is likely that they experience a catastrophic failure in processing, presumably resulting in relatively immediate and substantive disruption in the eye movement record. There are two aspects of our results that are important in relation to these specifications. First, if words are identified in parallel, then it seems likely that readers might detect a transposed word pair during fixations made on the word prior to the transposition. We obtained no evidence of any such disruption. Second, assuming more regressions for grammatical than ungrammatical sentences (due to processing associated with coherent syntactic formation), we might have expected to see more regressions from Target Words 1 and 2 for non-transposed than transposed sentences. Again, no such effects occurred. These aspects of the results, to us, do not appear to be consistent with the stipulations of SEAM. We did obtain two further results that can be considered consistent with SEAM, namely, that there were more regressions out of Target Word 2 and longer go-past reading times for Target Word 2 in the grammatically correct condition than when the sentences contained only an ungrammatical final word. This finding is consistent with the idea that readers make more regressions and spend more time re-reading sentences when constructing legal relative to illegal syntactic representations for them. We note that neither the E-Z Reader model, nor the OB1 Reader model can offer an explanation for these effects.

Whilst the SEAM model may offer an account of the regression out and go-past time results at Target Word 2 that we obtained for the grammatically correct sentences compared with sentences with a final ungrammatical word, we do consider that there may be an alternative explanation for these results. This is particualrly so, given that we did not see comparable effects when readers fixated the second word in the sentence (i.e., the Pre-Target word region) and the sentence contained a transposed word pair in the parafovea. In such a situation, readers would also have a word to the right of the fixated word that was ungrammatical. Despite the similarity in these circumstances, there was no evidence of any parafoveal ungrammaticality detection. The absence of similar effects earlier in the sentence led us to the view that the final word effects could have arisen particularly because the ungrammatical word was both sentence and line final. Recall that in eye movement experiments, researchers most often avoid placing critical regions at the end of sentences or in line final positions to ensure that readers have an opportunity to move their eyes rightward (in alphabetic languages like English) to fixate upcoming words and to avoid the possibility of experimental effects being clouded by sentence wrap-up and return sweep effects. Whilst we have no ready explanation for the final word effects that we observed, we suspect that the way in which readers process line final words, particularly when they are likely to regress to re-read earlier portions of a sentence, or make a return sweep, or make a large leftward saccade in order to prepare for the next trial of the experiment, may be different to the way in which the other words of the sentence are processed. For example, it may be the case that readers parafoveally process line final words to gain a sense of whether the upcoming word appears to be visually consistent with predictable candidates. If the upcoming word does look somewhat predictable, then it might be the case that a leftward eye movement is launched without a direct fixation to the line final word. In contrast, when the line final word does not resemble a predictable word (i.e., it is a word that is unlikely to appear in the sentence based on context, e.g., an ungrammatical word), then readers may be more likely to make a saccade to it in order to fixate it prior to making a subsequent leftward saccade. If this was the case, then it is likely that readers quite frequently fail to fixate line final words, and they do this to a greater degree than they skip line internal words. Clearly, further research is required to evaluate this suggestion, but regardless of whether it is, or is not, correct, it seems very likely that the effects at Target Word 2 that we observed here arose because the word that caused it was line final.

Before closing, there is a final aspect of the results that requires mention, namely, that response accuracy and response times with respect to grammaticality decisions were comparable between sentences containing a transposition alone (92 % accurate, 2261 ms) and sentences containing an ungrammatical final word only (91 % accurate, 2247 ms). In the present experiment, when participants read sentences containing a transposed word pair, they detected that transposition equally as often, and as rapidly, as they detected a sentence-final word that was grammatically illegal. According to the OB1 Reader model, this pattern of effects should not occur because processing in accord with the spatiotopic map should ensure that word pairs appearing out of order are accommodated and interpreted comfortably. Thus, the degree to which readers detect ungrammaticalities that arise due to the transposition of a pair of words in the sentence should be substantially reduced relative to their detection of single ungrammatical words that have been substituted within the sentence. Clearly, this was not the case in the present study. It seems very likely that detection rates and response times were comparable across these two conditions here because our stimuli were well pre-screened and tightly controlled to ensure that both types of manipulation produced sentences that were seen as ungrammatical, and participants were not required to make speeded judgments. To be explicit, the present results demonstrate very clearly that there is nothing particularly special about transposed word pairs in sentences in relation to them signalling, or not signalling, ungrammaticality.

6.1. Conclusions

In summary, in the present experiment we manipulated whether 5-word sentences did, or did not, contain a transposed word pair and did, or did not, contain a grammatically illegal final word. We measured participants' eye movements as they read and made grammaticality decisions to these stimuli. We found very high decision accuracy and relatively long times probably due to task requirements (reading for comprehension and unspeeded grammaticality decision) and tightly controlled experimental stimuli. Importantly, we obtained no disruption to processing prior to readers fixating the first word of the transposed word pair. Transposed words caused significant and rapid disruption to processing and an ungrammatical sentence final word attracted readers' fixations and caused increased re-reading. Overall, the results (arguably) are best accounted for by a serial eye movement model, though it is fair to say that none of the models of eye movement control adequately explains all aspects of the results. The patterns of effects we obtained, in the main, do seem to be consistent with those reported by Huang and Staub (2021a), and it is clearly the case that readers do detect transposed word pairs during reading when transpositions are taken to reflect an ungrammaticality and that these are detected incrementally, word by word, as sentences are read.

CRediT authorship contribution statement

Petar Atanasov: Conceptualization. **Simon P. Liversedge:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Conceptualization. **Federica Degno:** Writing – review & editing, Supervision, Software, Resources, Methodology, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to

influence the work reported in this paper.

Acknowledgments

We acknowledge support from a Ph.D. studentship from the University of Central Lancashire awarded to P.A., a Bial Foundation grant awarded to F.D. (201/20) and an ESRC grant (ES/R003386/1) awarded to S.P.L. We thank Chuanli Zang for helpful discussions during the write-up of this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cogpsych.2025.101715.

Data availability

Stimuli information, data, analyses scripts, trimming procedures for the generalised linear mixed effects models, software information, and tables with random effects estimates from the generalised linear mixed effects models for all measures are available at https://osf.io/uybpv/?view_only=371d7b96eac94be481a5378af74d1d10.

References

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of memory and language, 68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Clifton, C., Jr, Ferreira, F., Henderson, J. M., Inhoff, A. W., Liversedge, S. P., Reichle, E. D., & Schotter, E. R. (2016). Eye movements in reading and information processing: Keith Rayner's 40 year legacy. *Journal of Memory and Language, 86*, 1–19. https://doi.org/10.1016/j.jml.2015.07.004

Devescovi, A., Bates, E., D'amico, S., Hernandez, A., Marangolo, P., Pizzamiglio, L., & Razzano, C. (1997). An on-line study of grammaticality judgements in normal and aphasic speakers of Italian. *Aphasiology, 11*(6), 543–579. https://doi.org/10.1080/02687039708248490

Dufour, S., Mirault, J., & Grainger, J. (2022). Transposed-word effects in speeded grammatical decisions to sequences of spoken words. *Scientific Reports, 12*(1). https://doi.org/10.1038/s41598-022-26584-2

Engbert, R., Longtin, A., & Kliegl, R. (2002). A dynamical model of saccade generation in reading based on spatially distributed lexical processing. *Vision Research, 42*(5), 621–636. https://doi.org/10.1016/S0042-6989(01)00301-7

Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review, 112*(4), 777. https://doi.org/10.1037/0033-295X.112.4.777

Gibson, E., Piantadosi, S. T., Brink, K., Bergen, L., Lim, E., & Saxe, R. (2013). A noisy- channel account of crosslinguistic word-order variation. *Psychological science, 24*(7), 1079–1088. https://doi.org/10.1177/0956797612463705

Gomez, P., Ratcliff, R., & Perea, M. (2008). The overlap model: A model of letter position coding. *Psychological review, 115*(3), 577. https://doi.org/10.1037/a0012667

Hossain, J., & White, A. L. (2023). The transposed word effect is consistent with serial word recognition and varies with reading speed. *Cognition, 238*, Article 105512. https://doi.org/10.1016/j.cognition.2023.105512

Huang, K. J., & Staub, A. (2021a). Using eye tracking to investigate failure to notice word transpositions in reading. *Cognition, 216*, Article 104846. https://doi.org/10.1016/j.cognition.2021.104846

Huang, K. J., & Staub, A. (2021b). Why do readers fail to notice word transpositions, omissions, and repetitions? A review of recent evidence and theory. Language And Linguistics. *Compass, 15*(7). https://doi.org/10.1111/lnc3.12434

Huang, K. J., & Staub, A. (2022). Readers do not strongly rely on full-context information, but might utilize local word statistics, when 'correcting' word transposition errors in text. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (p. 44).

Huang, K. J., & Staub, A. (2023). The transposed-word effect does not require parallel word processing: Failure to notice transpositions with serial presentation of words. *Psychonomic Bulletin & Review, 30*, 393–400. https://doi.org/10.3758/s13423-022-02150-9

Johnson, R. L., Perea, M., & Rayner, K. (2007). Transposed-letter effects in reading: Evidence from eye movements and parafoveal preview. *Journal of Experimental Psychology: Human Perception and Performance, 33*(1), 209–229. https://doi.org/10.1037/0096-1523.33.1.209

Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science, 29*(3), 375–419. https://doi.org/10.1207/s15516709cog0000_25

Liu, Z., Li, Y., Cutter, M. G., Paterson, K. B., & Wang, J. (2022). A transposed-word effect across space and time: Evidence from Chinese. *Cognition, 218*, Article 104922. https://doi.org/10.1016/j.cognition.2021.104922

Liu, Z., Li, Y., Paterson, K. B., & Wang, J. (2020). A transposed-word effect in Chinese reading. *Attention, Perception, & Psychophysics, 82*(8), 3788–3794. https://doi.org/10.3758/s13414-020-02114-y

Liu, Z., Li, Y., & Wang, J. (2021). Context but not reading speed modulates transposed- word effects in Chinese reading. *Acta Psychologica, 215*, Article 103272. https://doi.org/10.1016/j.actpsy.2021.103272

Lo, S., & Andrews, S. (2015). To transform or not to transform: Using generalized linear mixed models to analyse reaction time data. *Frontiers in psychology, 6*, 1171. https://doi.org/10.3389/fpsyg.2015.01171

Lu, C. C., Bates, E., Li, P., Tzeng, O., Hung, D., Tsai, C. H., Lee, S., & Chung, Y. M. (2000). Judgements of grammaticality in aphasia: The special case of Chinese. *Aphasiology, 14*(10), 1021–1054. https://doi.org/10.1080/02687030050156593

Milledge, S. V., Bhatia, N., Mensah-Mcleod, L., Raghvani, P., McGowan, V. A., Elsherif, M. M., Cutter, M. G., Wang, J., Liu, Z., & Patterson, K. B. (2023). The transposed- word effect provides no unequivocal evidence for parallel processing. *Attention Perception Psychophysics*. https://doi.org/10.3758/s13414-023-02721-5

Mirault, J., Declerck, M., & Grainger, J. (2022). Fast priming of grammatical decisions: Repetition and transposed-word priming effects. *The Royal Society Open Science, 9*, Article 211082. https://doi.org/10.1098/rsos.211082

Mirault, J., Guerre-Genton, A., Dufau, S., & Grainger, J. (2020). Using virtual reality to study reading: An eye-tracking investigation of transposed-word effects. *Methods In Psychology, 3*, Article 100029. https://doi.org/10.1016/j.metip.2020.100029

Mirault, J., Leflaëc, C., & Grainger, J. (2022). Orthographic relatedness and transposed- word effects in the grammatical decision task. *Attention, Perception, &; Psychophysics, 84*(3), 1043–1051. https://doi.org/10.3758/s13414-021-02421-y
Mirault, J., Snell, J., & Grainger, J. (2018). You That Read Wrong Again! A Transposed- Word Effect in Grammaticality Judgments. *Psychological Science, 29*(12), 1922–1929. https://doi.org/10.1177/0956797618806296
Mirault, J., Vandendaele, A., Pegado, F., & Grainger, J. (2022). Transposed-word effects when reading serially. *PLOS ONE, 17*(11). https://doi.org/10.1371/journal.pone.0277116
Mirault, J., Vandendaele, A., Pegado, F., & Grainger, J. (2023). The impact of atypical text presentation on transposed-word effects. *Attention, Perception, & Psychophysics, 1–10*. https://doi.org/10.3758/s13414-023-02760-y
Pagán, A., Blythe, H. I., & Liversedge, S. P. (2021). The influence of children's reading ability on initial letter position encoding during a reading-like task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 47*(7), 1186–1203. https://doi.org/10.1037/xlm0000989
Pegado, F., & Grainger, J. (2019). Dissociating lexical and sublexical contributions to transposed-word effects. *Acta Psychologica, 201*, Article 102943. https://doi.org/10.1016/j.actpsy.2019.102943
Pegado, F., & Grainger, J. (2020a). A transposed-word effect in same-different judgments to sequences of words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 46*(7), 1364–1371. https://doi.org/10.1037/xlm0000776
Pegado, F., & Grainger, J. (2020b). On the noisy spatiotopic encoding of word positions during reading: Evidence from the change-detection task. *Psychonomic Bulletin & Review, 28*(1), 189–196. https://doi.org/10.3758/s13423-020-01819-3
Rabe, M. M., Paape, D., Mertzen, D., Vasishth, S., & Engbert, R. (2024). SEAM: An integrated activation-coupled model of sentence processing and eye movements in reading. *Journal of Memory and Language, 135*, Article 104496. https://doi.org/10.1016/j.jml.2023.104496
Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive. Psychology., 7*, 65–81. https://doi.org/10.1016/0010-0285(75)90005-5
Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology, 62*(8), 1457–1506. https://doi.org/10.1080/17470210902816461
Rayner, K., Angele, B., Schotter, E. R., & Bicknell, K. (2013). On the processing of canonical word order during eye fixations in reading: Do readers process transposed word previews? *Visual Cognition, 21*(3), 353–381. https://doi.org/10.1080/13506285.2013.791739
R Core Team. (2023). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.
Reichle, E. D. (2011). Serial-attention models of reading. In S. P. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), *The Oxford handbook of eye movements* (pp. 767–786). Oxford University Press.
Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Towards a model of eye movement control in reading. *Psychological. Review., 105*(1), 125–157. https://doi.org/10.1037/0033-295X.105.1.125
Reichle, E. D., & Sheridan, H. (2015). E-Z Reader: An overview of the model and two recent applications. In A. Pollatsek, & R. Treiman (Eds.), *The Oxford handbook of reading* (pp. 277–290). Oxford University Press.
Snell, J., & Grainger, J. (2019a). Readers are parallel processors. *Trends in Cognitive Sciences, 23*(7), 537–546. https://doi.org/10.1016/j.tics.2019.04.006
Snell, J., & Grainger, J. (2019b). Word position coding in reading is noisy. *Psychonomic Bulletin & Review, 26*(2), 609–615. https://doi.org/10.3758/s13423-019-01574-0
Snell, J., & Melo, A. N. (2024). Do Love You Me? Failure to Notice Word Transpositions is Induced by Parallel Word Processing. Journal of. *Cognition, 7*(1). https://doi.org/10.5334/joc.335
Snell, J., Mirault, J., Theeuwes, J., & Grainger, J. (2023). Readers use word length information to determine word order. *Journal of Experimental Psychology: Human Perception and Performance, 49*(6), 753–758. https://doi.org/10.1037/xhp0001107
Snell, J., van Leipsig, S., Grainger, J., & Meeter, M. (2018). OB1-reader: A model of word recognition and eye movements in text reading. *Psychological Review, 125*(6), 969–984. https://doi.org/10.1037/rev0000119
Spinelli, G., Yang, H., & Lupker, S. J. (2024). Rapid Serial Visual Presentation of transposed-word sequences in the grammatical decision task: An examination of the roles of temporal and spatial cues to word order. *Language, Cognition and Neuroscience, 39*(6), 773–792. https://doi.org/10.1080/23273798.2024.2360168
Tiffin-Richards, S. P. (2024). The effect of word transpositions on grammaticality judgements in first and second language sentence reading. *Quarterly Journal of Experimental Psychology, 77*(1), 204–216. https://doi.org/10.1177/17470218231161433
Van Heuven, W. J., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *Quarterly journal of experimental psychology, 67*(6), 1176–1190. https://doi.org/10.1080/17470218.2013.850521
Veldre, A., Drieghe, D., & Andrews, S. (2017). Spelling ability selectively predicts the magnitude of disruption in unspaced text reading. *Journal of Experimental Psychology: Human Perception and Performance, 43*(9), 1612–1628. https://doi.org/10.1037/xhp0000425
Venables, W. N., & Ripley, B. D. (2002). *Modern Applied Statistics with S* (Fourth edition). New York: Springer. https://www.stats.ox.ac.uk/pub/MASS4/.
Von der Malsburg, T., & Angele, B. (2017). False positives and other statistical errors in standard analyses of eye movements in reading. *Journal of memory and language, 94*, 119–133. https://doi.org/10.1016/j.jml.2016.10.003
Warren, T., White, S. J., & Reichle, E. D. (2009). Investigating the causes of wrap-up effects: Evidence from eye movements and E–Z reader. *Cognition, 111*(1), 132–137. https://doi.org/10.1016/j.cognition.2008.12.011
Wen, Y., Mirault, J., & Grainger, J. (2021a). Fast syntax in the brain: Electrophysiological evidence from the rapid parallel visual presentation paradigm (RPVP). *Journal of Experimental Psychology: Learning, Memory, and Cognition, 47*(1), 99–112. https://doi.org/10.1037/xlm0000811
Wen, Y., Mirault, J., & Grainger, J. (2021b). The transposed-word effect revisited: The role of syntax in word position coding. *Language, Cognition and Neuroscience, 36*(5), 668–673. https://doi.org/10.1080/23273798.2021.1880608
Wen, Y., Mirault, J., & Grainger, J. (2022). A transposed-word effect on word-in- sequence identification. *Psychonomic Bulletin & Review, 29*(6), 2284–2292. https://doi.org/10.3758/s13423-022-02132-x
Wen, Y., Mirault, J., & Grainger, J. (2024). On relative word length and transposed-word effects. *Journal of Experimental Psychology: Human Perception and Performance, 50*(9), 934. https://doi.org/10.1037/xhp0001229
Wen, Y., Snell, J., & Grainger, J. (2019). Parallel, cascaded, interactive processing of words during sentence reading. *Cognition, 189*, 221–226. https://doi.org/10.1016/j.cognition.2019.04.013
Westfal, J. (2016). PANGEA: Power Analysis for General Anova designs. http://jakewestfall.org/publications/pangea.pdf.